

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 861 663 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication: 02.09,1998 Bulletin 1998/36

(51) Int CI.⁶ **A61K 38/20** // C07K14/54

(21) Application number: 98301352.5

(22) Date of filing: 24.02.1998

(84) Designated Contracting States:

AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC

NL PT SE

Designated Extension States:

AL LT LV MK RO SI

(30) Priority: 25.02.1997 JP 55468/97

(71) Applicant: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO Okayama-shi Okayama (JP)

(72) Inventors:

Gillespie, Matthew Todd,
 Department of Medicine
 Med. Res 41 Victoria Parade Fitzroy 3065 (AU)

 Horwood, Nicole Joy, Department of Medicine Med. Res 41 Victoria Parade Fitzroy 3065 (AU)

 Udagawa, Nobuyuki Kashiwa-shi, Chiba (JP)

Kurimoto, Masashi
 Okayama-shi, Okayama (JP)

 (74) Representative: Daniels, Jeffrey Nicholas et al Page White & Farrer
 54 Doughty Street
 London WC1N 2LS (GB)

Remarks:

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Osteoclastgenic inhibitory agent comprising interleukin-18

(57) An osteoclastgenic inhibitory agent which comprises an interleukin-18 and/or its functional equivalent. The agent can be arbitrarily used as an ingredient for

cell culture and agents for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Description

5

10

15

20

25

35

40

45

50

55

The present invention relates to an osteoclastgenic inhibitory agent comprising an interleukin-18 (hereinafter abbreviated as "IL-18") or its functional equivalent.

Osteoblasts' bone formation and osteoclasts' bone resorption are well balanced in healthy living bodies, and this keeps the bone tissues in normal conditions while old bone tissues are being replaced with fresh ones without altering the original bone shape. The phenomenon plays an important role in keeping living bodies' homeostasis such as the controlling of blood calcium concentration within a desired range. Once the balance is lost, especially when the bone resorption level exceeds the bone formation level, bone-related diseases and other diseases may be induced. Therefore, elucidation of the whole mechanism of bone resorption in living bodies, particularly, elucidation of osteoclasts is greatly highlighted due to scientific and clinical significance thereof.

However, the mechanism of osteoclast formation has not yet been completely elucidated even though interleukin 1 as a promoter and interleukin 4 as an inhibitor were found. This is because, similarly as various phenomena in living bodies, osteoclast formation in living bodies is controlled by the close and complicated relationship between promoters and inhibitors. Based on these, it is greatly expected to establish an effective osteoclastgenic inhibitory agent from the viewpoint of scientific and clinical aspects.

The object of the present invention is to provide a novel and effective osteoclastgenic inhibitory agent. To solve the object the present inventors energetically studied for IL-18, i.e., one of cytokines as communication transferring substances in immune systems, which induces production of interferon-y (hereinafter abbreviated as "IFN-γ"), an important biologically active substance for immunocompetent cells, and granulocyte/macrophage colony-stimulating factor (hereinafter abbreviated as "GM-CSF"), and augments cytotoxicity and induces formation of killer cells. At the finding, IL-18 was described as an interferon-γ-inducing factor as reported by Haruki OKAMURA in Japanese Patent Kokai Nos. 27,189/96 and 193,098/96, and in *Nature*, Vol. 378, No. 6,552, pp. 88-91 (1995), and then called IL-18 according to the proposal of Shimpei USHIO et al., in *The Journal of Immunology*, Vol. 156, pp. 4,274-4,279 (1996).

The present inventors found that a particular gene, capable of inhibiting osteoclast formation from osteoclastic precursor cells *in vitro*, is specifically expressed in quantities in stroma cells derived from mouse myeloma. Their further detailed analysis revealed that (i) the gene encodes IL-18 that includes SEQ ID NO: 7 as a core sequence, (ii) IL-18 and functional equivalents thereof effectively inhibit osteoclast formation, and (iii) the inhibition is mainly due to the action of GM-CSF induced and produced by IL-18.

Based on these, the present inventors solved the present object by an osteoclastgenic inhibitory agent comprising IL-18 or its functional equivalent as an effective ingredient.

FIG. 1 shows the structure of the recombinant DNA pKGFHH2.

FIG. 2 shows the structure of the recombinant DNA pCSHIGIF/MUT35.

FIG. 3 shows the structure of the recombinant DNA pCSHIGIF/MUT42.

FIG. 4 shows the structure of the recombinant DNA pBGHuGF.

FIG. 5 shows the structure of the recombinant DNA pKGFMH2.

In these figures, KGFHH2 cDNA means a cDNA encoding the IL-18 according to the present invention: IGIF/MUT35; a DNA encoding the IL-18 according to the present invention: IGIF/MUT42; a DNA encoding the IL-18 according to the present invention: HuIGIF; a chromosomal DNA encoding the IL-18 according to the present invention: KGFMH2 cDNA; a cDNA encoding the IL-18 according to the present invention: 5S; a gene for 5S ribosomal RNA: Ptac; a tac promoter: rrnBT1T2; a termination region of a ribosomal RNA operon: AmpR; an ampicillin resistent gene: pBR322ori; a replication origin of *Escherichia coli*: CMV; a cytomegalovirus promoter: IFNss; a nucleotide sequence encoding a signal peptide for subtype α2b of human interferon-α.

The present invention relates to an osteoclastgenic inhibitory agent comprising IL-18 or its functional equivalent as an effective ingredient. The wording "IL-18" as referred to in the invention includes polypeptides with the above property independently of their sources and origins. For example, the IL-18 used in the present invention includes, as internal partial amino acid sequences, the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3, as well as SEQ ID NO: 4 and SEQ ID NO: 5, and includes the amino acid sequence of SEQ ID NO: 6 or SEQ ID NO: 7 as a whole. The wording "functional equivalent(s)" as referred to in the present invention includes (i) those wherein one or more amino acids in the amino acid sequence of IL-18 are replaced with different amino acids, (ii) those wherein one or more amino acids are added to the N- and/or C-termini of the amino acid sequence of IL-18, (iii) those wherein one or more amino acids are inserted into the internal sites of the amino acid sequence of IL-18 are deleted, and (v) those wherein one or more amino acids in the N- and/or C-terminal regions of the amino acid sequence of IL-18 are deleted; all of these modifications should be made within the range that does not substantially lose the property of osteoclast formation by IL-18 among the inherent property of IL-18. Examples of such functional equivalents are described along with their detailed amino acid sequences in Japanese Patent Application No. 20,906/97 by the same applicant of the present applicant, i.e., polypeptides which are capable of inducing production of interferon-gamma by immunocompe-

tent cells, wherein said polypeptides contain either amino acid sequence wherein one or more cysteines are replaced with different amino acid(s) while leaving respective consensus sequences as shown in SEQ ID NOs; 1, 2 and 4 intact, or that wherein one or more amino acids are added, removed and/or replaced at one or more sites including those in the consensus sequences but excluding those of the replaced cysteine. The different amino acids to replace the cysteine(s) are not restricted to any types, as far as the resulting polypeptide, containing an amino acid sequence replaced with the different amino acid(s), exhibits an activity of inducing production of IFN-γ by immunocompetent cells in the presence or absence of an appropriate cofactor, as the wild-type polypeptides containing SEQ ID NOs: 1, 2 and 4 as consensus partial amino acid sequences, and a stability significantly higher than that of the wild-type polypeptides. The different amino acids include serine, threonine, alanine, valine, leucine, isoleucine, histidine, tyrosine, phenylalanine, tryptophan, and methionine, among which the most preferable amino acid is serine or alanine. Embodiments of the amino acid sequences, containing SEQ ID NOs: 1, 2 and 4 as consensus partial amino acid sequences, in which one or more cysteines are to be replaced with different amino acid(s) are the wild-type polypeptides containing SEQ ID NO: 6 or 7. SEQ ID NO: 6 contains cysteines at the 38th, 68th, 76th, and 127th positions from the N-terminus, SEQ ID NO: 7 contains cysteines at the 7th, 75th, and 125th positions. The polypeptides include those containing the amino acid sequence of any one of SEQID NOs: 20-26, which are derived from the wild-type polypeptide containing SEQID NO: 6, those containing the amino acid sequence of SEQ ID NO: 27 or 28, which are derived from the wild-type polypeptide containing the amino acid sequence of SEQ ID NO: 7, and those containing an amino acid sequence derived from any one of SEQ ID NOs: 20-28 by adding, removing, and/or replacing one or more amino acids to and/ or at position(s) excepting the positions where the cysteine(s) have been replaced while retaining the desired biological activities and stability. The wording "one or more amino acids" means the number of amino acids which conventional methods such as site-directed mutagenesis can usually add, remove or replace. The polypeptides containing any one of SEQ ID NOs: 20-28 possess both stability and biological activities significantly higher than those of the wild-type

15

20

25

30

40

The functional equivalents as referred to in the present invention further include glycosylated polypeptides of IL-18 and the above polypeptides. Any of these IL-18 and functional equivalents thereof, both of which are included to and referred to as "IL-18" in the present invention, unless specified otherwise, can be used in the present invention independently of their origins; those prepared by separating from natural sources such as cell cultures and from artificially synthesized ones using recombinant DNA technology and peptide synthesis.

With economical viewpoint, methods of recombinant DNA technology are advantageously used: generally, desired IL-18 can be obtained by introducing DNAs encoding IL-18 into appropriate hosts derived from microorganisms, plants, and animals to form transformants, culturing the transformants in nutrient culture media in a conventional manner, and purifying the cultures by conventional methods used for purifying cytokines. Any DNAs can be used as the above DNAs as long as they contain a DNA encoding IL-18, and can be suitably selected depending on the purpose of the use of the present osteoclastgenic inhibitory agent or on the recombinant DNA technology used. For example, Japanese Patent Kokai Nos. 193,098/96, 231,598/96, and 27,189/96 by the same applicant of the present invention disclose in detail methods for producing IL-18 by culturing transformed microorganisms into which DNAs including a cDNA encoding mouse or human IL-18 are introduced; and Japanese Patent Application No. 185,305/96 by the same applicant of the present invention discloses in detail a method for producing IL-18 encoding human IL-18 by culturing transformed animal cells which have an introduced DNA that includes a chromosomal DNA encodes human IL-18. Japanese Patent Application No. 20,906/97 by the same applicant of the present invention discloses in detail a method for producing IL-18 by culturing transformed animal cells having an introduced DNA which includes a DNA encoding a functional equivalent of human IL-18.

The aforesaid recombinant DNA technology has an economical advantage, but depending on the hosts and DNA sequences used, the IL-18 thus obtained may have somewhat different physicochemical property from those of IL-18 produced and functions *in vivo*. Japanese Patent Application No. 67,434/96 by the same applicant of the present invention discloses in detail a preparation of IL-18 using established human cell lines as natural sources, and Japanese Patent Application No. 213,267/96 by the same applicant also discloses in detail the preparation using an interleukin-1β-converting enzyme. The IL-18 obtained by those preparations can be estimated to have substantially the same or equal physicochemical property to that of IL-18 that is produced and functions in *vivo*, and the yield can be estimated to be slightly lower. However, such IL-18 has an advantage that it has a fewer side effects when used as pharmaceuticals directed to administering to warm-blooded animals in general and including humans. When applying purification methods using monoclonal antibodies specific to IL-18, as disclosed in Japanese Patent Application No. 231,598/96 by the same applicant of the present invention, a relatively-high purity IL-18 can be obtained in a minimum labor and cost.

The present osteoclastgenic inhibitory agent comprising the aforesaid IL-18 includes any types and forms usable to inhibit osteoclast formation both in *vivo* and in *vitro*. The present agent can be advantageously used as ingredients for cell culture media for animal cells, which satisfactorily inhibit osteoclast formation, maintain, proliferate, and/or differentiate the desired cells; components of screening kits for bone-related therapeutic agents; bone-resorption regulatory agents; and agents for osteoclast-related diseases. The bone-resorption regulatory agents include medica-

ments and health foods that exert an osteoclastgenic inhibitory activity in *vivo*, control bone resorption to normal conditions, and improve unfavorable physical conditions such as a relatively-insignificant arthralgia. The agents for osteoclast-related diseases include medicaments used to prevent and/or treat diseases caused by an excessive osteoclast formation and/or its function. Examples of such diseases are hypercalcemia, osteoclastoma. Behcet's syndrome, osteosarcoma, arthropathy, chronic rheumatoid arthritis, deformity ostitis, primary hyperthyroidism, osteopenia, and osteoporosis. Varying depending on the types of agents and diseases to be treated, the present agent is usually formulated into a liquid, paste, or solid form which contains 0.000002-100 w/w %, preferably, 0.0002-0.5 w/w % of IL-18.

The present osteoclastgenic inhibitory agent can be IL-18 alone or compositions comprising IL-18 and one or more other ingredients such as carriers, excipients, diluents, adjuvants, antibiotics, and proteins such as serum albumin and gelatin as stabilizers; saccharides such as glucose, maltose, maltotriose, maltotetraose, trehalose, sucrose, isomaltose, lactose, panose, erlose, palatinose, lactosucrose, raffinose, fructooligosaccharide, galactooligosaccharide, lentinan, dextrin, pullulan, and sugar alcohols including sorbitol, maltitol, lactitol, and maltotriitol; buffers comprising phosphates or citrates mainly; and reductants such as 2-mercaptoethanol, dithiothreitol, and reduced glutathione; and optionally biologically active substances such as interferon-a, interferon-p, interferon-γ, interleukin-2, interleukin-3, interleukin-6, interleukin-12, TNF-α, TNF-β, GM-CSF, estrogen, progesterone, chlormadinone acetate, calcitonin, somatokine, somatomedin, insulin-like growth factor, ipriflavone, parathyroid hormone (PTH), norethisterone, busulfan, ancitabine, cytarabine, fluorouracil, tetrahydrofurfuryl fluorouracil, methotrexate, vitamin D₂, active vitamin D, Krestin® or polysaccharide K, L-asparaginase, and OK-432 or Picibanil®; and calcium salts such as calcium lactate, calcium chloride, calcium monohydrogenphosphate, and L-calcium L-aspartate. When used as agents for administering to warmblooded animals in general and including humans, i.e., agents for osteoclast-related diseases, the present agent can be preferably formulated into compositions by appropriately combining with one or more of the above physiologically-acceptable substances.

The present osteoclastgenic inhibitory agent includes medicaments in a unit dose form used for administering to warm-blooded animals in general and including humans. The wording "unit dose form" means those which contain IL-18 in an amount suitable for a daily dose or in an amount up to four fold by integers or up to 1/40 fold of the dose, and those in a physically separated and formulated form suitable for prescribed administrations. Examples of such formulations are injections, liquids, powders, granules, tablets, capsules, troches, collyriums, nebulas, and suppositories.

The present agent as an osteoclastgenic inhibitory agent effectively treat and prevent osteoclast-related diseases independently of oral and parenteral administrations. Varying depending on the types and symptoms of patients' diseases, the present agent can be administered to the patients orally, intradermally, subcutaneously, muscularly, or intravenously at a dose of about $0.5~\mu g$ to 100~mg per shot, preferably, at a dose of about $2~\mu g$ to 10~mg per shot of IL-18, 2-6 fold a day or 2-10 fold a week for one day to one year.

In the below, with reference to experiments, the preparation, physicochemical property, and biological activity of the IL-18 according to the present invention are described:

Experiment 1

10

20

25

35

40

45

50

Preparation of human IL-18

According to the method in Japanese Patent Kokai No. 231,598/96 by the same applicant of the present invention, an autonomously-replicable recombinant DNA, pKGFHH2, linked to a cDNA encoding human IL-18, was prepared. Dideoxyribonucleotide sequencing analyzed that, as shown in FIG. 1, in the recombinant DNA, KGFHH2 cDNA containing the base sequence of SEQ ID NO: 8 was linked to the downstream of Ptac, a Tac promoter. The recombinant DNA pKGFHH2 contained the amino acid sequences of SEQ ID NOs: 1 to 5; these amino acid sequences were respectively encoded by nucleotides 46-63, 88-105, 400-420, 151-165, and 214-228 in SEQ ID NO: 8.

According to the method in Japanese Patent Kokai No. 231,598/96, the recombinant DNA pKGFHH2 was introduced into an *Escherichia coli* Y1090 strain, ATCC 37197, and the strain was cultured. The produced polypeptide was purified by immunoaffinity chromatography to obtain a purified human IL-18 with a purity of at least 95% in a yield of about 25 mg/*ℓ* culture. According to the method in Japanese Patent Kokai No. 193,098/96 by the same applicant of the present invention, the purified human IL-18 was analyzed for biological activity and physicochemical property as indicated below: When culturing human lymphocytes, collected by a conventional manner from a healthy donor, in the presence of the purified human IL-18, IFN-γ production was observed depending on the concentration of IL-18, resulting in a confirmation that IL-18 has an activity of inducing IFN-γ production by lymphocytes as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in *Nature*, Vol. 227, pp. 680-685 (1970), the purified IL-18 was subjected to SDS-PAGE, resulting in a major band with an IFN-γ inducing activity at a position corresponding to 18,500±3,000 daltons. The IL-18 gave a pl of 4.9±1.0 as determined by conventional chromatofocusing. Conventional analysis using "PROTEIN SEQUENCER MODEL 473A", an apparatus of Applied Biosystems, Inc., Foster City,

USA revealed that the IL-18 had the amino acid sequence of SEQ ID NO. 9, i.e., the amino acid sequence of SEQ ID NO: 8 where a methionine residue was linked to the N-terminus.

Experiment 2

5

10

15

25

30

35

40

45

50

55

Preparation of human IL-18

According to the method in Japanese Patent Application No. 67.434/96 by the same applicant of the present invention, THP-1 cells, ATCC TIB 202, a human monocyte cell line derived from a male with acute monocytic leukemia, were inoculated to the dorsum subcutaneous tissues of new born hamsters, followed by feeding the hamsters for three weeks. Tumor masses, about 15 g weight each, formed in the subcutaneous tissues of each hamster, were extracted, dispersed in media, and disrupted. The polypeptide obtained from the disrupted cells was purified by immunoaffinity chromatography to obtain a purified human IL-18 in a yield of an about 50 ng/head.

Similarly, according to the method in Japanese Patent Application No. 67,434/96, the purified human IL-18 was analyzed and determined for biological activity and physicochemical property as indicated below: It was confirmed that culturing human lymphocytes, collected from healthy donors in a conventional manner, in the presence of different concentrations of the human IL-18, resulted in an IL-18 dose-dependent IFN-y production. This revealed that the human IL-18 has a biological activity of inducing IFN-γ production by lymphocytes as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in Nature, Vol. 227, pp. 680-685 (1970), the purified human IL-18 was subjected to SDS-PAGE using 2 w/v % dithiothreitol as a reductant, resulting in a major band with an IFN-y production inducing activity at a position corresponding to 18,000-19,500 daltons. According to the peptide map disclosed in Japanese Patent Application No. 67,434/96, the human IL-18 was treated with clostripain commercialized by Sigma Chemical Company, Missouri, USA, to obtain polypeptide fragments, followed by subjecting the fragments for fractionation to high-performance liquid chromatography (HPLC) using "ODS-120T", a column commercialized by Tosoh Corporation, Tokyo, Japan, and analyzing the amino acid sequences of the fragments from the N-terminus to reveal the following amino acid sequences of SEQ ID NOs: 10 to 13. These amino acid sequences were completely coincided with amino acids 148-157, 1-13, 45-58, and 80-96 in SEQ ID NO: 6. The data shows that the human IL-18 obtained in Experiment 2 has the amino acid sequence of SEQ ID NO: 6 and all the partial amino acid sequences of SEQ ID NOs: 1 to 5.

Experiment 3

Preparation of functional equivalents

According to the method in Japanese Patent Application No. 20,906/97 by the same applicant of the present invention, it was prepared an autonomously-replicable recombinant DNA, pCSHIGIF/MUT35, was linked to a DNA encoding a functional equivalent of human IL-18 where cysteines 38, 68, and 76 in SEQ ID NO: 6 were respectively replaced with serine, serine, and alanine. Dideoxyribonucleotide sequence analysis revealed that as shown in FIG. 2, in the recombinant DNA, DNA IGIF/MUT35 with SEQ ID NO: 14 linked to the downstream of a base sequence encoding a signal peptide of subtype α2b in human interferon-a in the same reading-frame, as reported by K. Henco et al., in *Journal of Molecular Biology,* Vol. 185, pp. 227-260 (1985), and had a stop codon for protein synthesis at further downstream. As shown in parallel in SEQ ID NO: 14, the amino acid sequence encoded by the recombinant DNA corresponded to SEQ ID NO: 6 where cysteines 38, 68, and 76 in SEQ ID NO: 6 were respectively replaced with serine, serine, and alanine. The recombinant DNA contained a nucleotide which encodes all the amino acid sequences of SEQ ID NOs: 1 to 4 and the one of SEQ ID NO: 5 where cysteine at amino acid 5 in SEQ ID NO: 5 was replaced with alanine. These amino acid sequences were respectively encoded by nucleotides 46-63, 88-105, 400-420, 151-165, and 214-228 in SEQ ID NO: 14.

According to the method in Japanese Patent Application No. 20,906/97 by the same applicant of the present invention, the recombinant DNA pCSHIGIF/MUT35 was introduced into COS-1 cells, ATCC CRL 1650, an established cell line derived from SV40 transformed African Green monkey kidney, followed by culturing the transformed cells. The produced polypeptide in the culture was purified by immunoaffinity chromatography to obtain a purified functional equivalent of human IL-18 in a yield of about 40 ng/ml culture. According to the method in Japanese Patent Application No. 20,906/97, the purified functional equivalent was analyzed and determined for biological activity and physicochemical property as indicated below: When culturing KG-1 cells, ATCC CCL 246, an established cell line derived from human acute myelogenous leukemia, in the presence of different concentrations of the purified functional equivalent of human IL-18, IFN-γ production was observed depending on the concentration of the IL-18, revealing that the IL-18 has a biological activity of inducing IFN-γ production by KG-1 cells as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in *Nature*, Vol. 227, pp. 680-685 (1970), the purified functional equivalent was

subjected to SDS-PAGE in the presence of 2 w/v % dithiothreitol as a reductant, resulting in a major band with an IFN-yproduction inducing activity at a position corresponding to 18,000-19,500 daltons. Conventional analysis using "PRO-TEIN SEQUENCER MODEL 473A", an apparatus of Applied Biosystems, Inc., Foster City, USA, revealed that the N-terminal region of the functional equivalent had the amino acid sequence of SEQ ID NO: 15 which corresponded to the amino acid sequence in the N-terminal region as shown in parallel in SEQ ID NO: 14.

Experiment 4

5

10

15

20

25

30

35

40

45

50

55

Preparation of functional equivalent

According to the method in Japanese Patent Application No. 20.906/97 by the same applicant of the present invention, it was prepared an autonomously-replicable recombinant DNA, pCSHIGIF/MUT42, which was linked to a DNA encoding for a functional equivalent of human IL-18 where cysteines 38, 68, 76, and 127 in SEQ ID NO; 6 were respectively replaced with serine, serine, alanine, and serine. Dideoxyribonucleotide sequencing revealed that, as shown in FIG. 3, in the recombinant DNA, DNA IGIF/MUT42 with SEQ ID NO; 16 linked to the downstream of a base sequence encoding a signal peptide for subtype α 2b of human interferon-a in the same reading frame, as reported by K. Henco et al., in *Journal of Molecular Biology*, Vol. 185, pp. 227-260 (1985), and had a stop codon for protein synthesis at further downstream. As shown in parallel in SEQ ID NO; 16, the amino acid sequence encoded by the recombinant DNA corresponded to SEQ ID NO; 6 where cysteines 38, 68, 76, and 127 in SEQ ID NO; 6 were respectively replaced with serine, serine, alanine, and serine. The recombinant DNA contained a nucleotide sequence which encodes all the amino acid sequences of SEQ ID NOs; 1 to 4 and the one of SEQ ID NO; 5 where cysteine 5 in SEQ ID NO; 5 was replaced with alanine. These amino acid sequences were respectively encoded by nucleotides 46-63, 88-105, 400-420, 151-165, and 214-228 in SEQ ID NO; 16.

According to the method in Japanese Patent Application No. 20,906/97 by the same applicant of the present invention, the recombinant DNA pCSHIGIF/MUT42 was introduced into COS-1 cells, followed by culturing the cells. The produced polypeptide in the culture was purified by immunoaffinity chromatography to obtain a purified functional equivalent of human IL-18 in a yield of about 20 ng/ml culture. According to the method in Japanese Patent Application No. 20,906/97, the purified functional equivalent was analyzed and determined for biological activity and physicochemical property as indicated below: When cultured KG-1 cells in the presence of different concentrations of the purified functional equivalent, a dose-dependent IFN-γ production was observed, and this revealed that the functional equivalent has a biological activity of inducing IFN-γ production by KG-1 cells as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in *Nature*, Vol. 227, pp. 680-685 (1970), the purified functional equivalent was subjected to SDS-PAGE in the presence of 2 w/v % dithiothreitol as a reductant, resulting in a major band with an IFN-γ inducing activity at a position corresponding to 18,000-19,500 daltons. Conventional analysis using "PROTEIN SEQUENCER MODEL 473A*, an apparatus of Applied Biosystems, Inc., Foster City, USA, revealed that the N-terminal region of the functional equivalent had the amino acid sequence of SEQ ID NO: 15 which completely corresponded to the amino acid sequence in the N-terminal region as shown in parallel in SEQ ID NO: 16.

Experiment 5

Preparation of human IL-18

According to the method in Japanese Patent Application No. 185,305/96 by the same applicant of the present invention, an autonomously-replicable recombinant DNA, pBGHuGF, linked to a chromosomal DNA encoding human IL-18, was obtained. Dideoxyribonucleotide sequencing analysis revealed that as shown in FIG. 4, in the recombinant DNA, a chromosomal DNA, which encodes human IL-18, i.e., DNA HulGIF with SEQ ID NO: 17, was linked to the downstream of a restriction site by a restriction enzyme, *Hind* III. As shown in SEQ ID NO: 17, the chromosomal DNA HulGIF consists of 11,464 bp where the exon was fragmented by four introns positioning at nucleotides 83-1,453, 1,466-4,848, 4,984-6,317, and 6,452-11,224. Among the resting nucleotide sequence excluding these introns, nucleotides 3-11,443 from the 5'-terminus are the part that encodes a precursor of human IL-18, and nucleotides 4,866-4,983 are the part that encodes an active human IL-18. The chromosomal DNA contained nucleotides sequences encoding SEQ ID NOs: 1 to 5; these amino acid sequences were respectively encoded by nucleotides 4,911-4,928, 4,953-4,970, 11,372-11,392, 6,350-6,364, and 6,413-6,427 in SEQ ID NO: 17.

According to the method in Japanese Patent Application No. 185,305/96, the recombinant DNA pBGHuGF was introduced into CHO-K1 cells, ATCC CCL 61, an established cell line derived from Chinese hamster ovary, followed by culturing the cells. The culture supernatant was contacted with a supernatant of cell disruptant prepared from a THP-1 cell culture to produce a polypeptide which was then purified by immunoaffinity chromatography to obtain a purified human IL-18 in a yield of about 15 mg/ ℓ culture. According to the method in Japanese Patent Application No.

135.305/96. the polypeptide was analyzed and determined for biological activity and physicochemical property as indicated below: It was confirmed that human lymphocytes, which were collected from a healthy donor, produced IFN-7 depending on the purified human IL-18 concentration when cultured at different concentrations of the human IL-18, revealing that the human IL-18 has a biological activity of inducing IFN-7 production by lymphocytes as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in *Nature*, Vol. 227, pp. 680-685 (1970), the purified human IL-18 was subjected to SDS-PAGE in the presence of 2 w/v % dithiothreitol as a reductant, resulting in a major band with an IFN-7 inducing activity at a position corresponding to 18.000-19,500 daltons. The N-terminal region of the human IL-18 contained the amino acid sequence of SEQ ID NO: 15 which completely corresponded to the amino acid sequence in the N-terminal region of SEQ ID NO: 17 for an active IL-18. Experiment 6

Preparation of mouse IL-18

10

15

20

25

30

40

50

To a 0.5-ml reaction tube were added 8 µl of 25 mM magnesium chloride. 10 µl of 10 x PCR buffer, one µl of 25 mM dNTP mix, one µl of 2.5 units/µl of amplitaq DNA polymerase, one ng of a recombinant DNA, which encodes mouse IL-18 having the nucleotide sequence of SEQ ID NO: 18 and the amino acid sequence of SEQ ID NO: 7, prepared from a phage DNA clone according to the method in Japanese Patent Kokai No. 27,189/96, and adequate amounts of a sense and antisense primers having nucleotide sequences represented by 5'-ATAGAATTCAAAT-GAACTTTGGCCGACTTCACTG-3' and 5'-ATAAAGCTTCTAACTTTGATGTAAGTT-3', respectively, which were chemically synthesized based on the amino acid sequences nearness to the N- and C-termini of SEQ ID NO: 7, and the mixture solution was brought up to a volume of 100 µl with sterilized distilled water. The solution thus obtained was subjected in a usual manner to PCR reaction of the following three cycles of successive incubations at 94°C for one minute, 43°C for one minute, and 72°C for one minute.

The product obtained by the PCR reaction and "pCR-Script SK (+)", a plasmid vector commercialized by Stratagene Cloning Systems, California, USA, were in a conventional manner ligated together using a DNA ligase into a recombinant DNA which was then introduced into "XL-1 Blue MRF'Kan", an *Escherichia coli* strain commercialized by Stratagene Cloning Systems, California, USA., to obtain a transformant. The transformant was inoculated to L-broth (pH 7.2) containing 50 µg/ml ampicillin, followed by the incubation at 37°C for 18 hours under shaking conditions. The culture was centrifuged to obtain the proliferated transformants which were then treated with a conventional alkali-SDS method to isolate a recombinant DNA. A portion of the recombinant DNA isolated was analyzed by dideoxyribonucle-otide sequencing, revealing that the recombinant DNA contained restriction sites of *Eco* RI and *Hind* III at the 5'- and 3'-termini of SEQ ID NO: 18, respectively, and a DNA containing a methionine codon for initiating polypeptide synthesis and a TAG codon for terminating polypeptide synthesis, which were located in just before and after the N- and C-termini of the amino acid sequence as shown in parallel in SEQ ID NO: 18. The recombinant DNA contained the nucleotide sequences of SEQ ID NOs: 1 to 5. These amino acid sequences were encoded by nucleotides 46-63, 85-102, 394-414, 148-162, and 211-225 in SEQ ID NO: 18.

The remaining portion of the recombinant DNA was in a conventional manner cleaved with restriction enzymes of *Eco* RI and *Hind* II, and the resulting 0.1 µg of an *Eco* RI-Hind III DNA fragments, obtained by using "DNA LIGATION KIT VER 2", a DNA ligation kit commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, and 10 ng of pKK223-3, a plasmid vector commercialized by Pharmacia LKB Biotechnology AB, Uppsala, Sweden, which had been cleaved with a restriction enzyme were linked together, by incubating at 16°C for 30 min to obtain an autonomously-replicable recombinant DNA, pKGFMH2. Using competent cell method, an *Escherichia coli* Y1090 strain, ATCC 37197, was transformed using the recombinant DNA pKGFMH2, and the resulting transformant, KGFMH2, was inoculated to L-broth (pH 7.2) containing 50 µg/ml ampicillin, and cultured at 37°C for 18 hours under shaking conditions. The culture was centrifuged to collect the proliferated transformants, followed by applying a conventional SDS-alkali method to a portion of the transformants to extract the recombinant DNA pKGFMH2. Dideoxyribonucleotide sequencing analysis revealed that, as shown in FIG. 5, KGFMH2 cDNA containing the nucleotide sequence of SEQ ID NO: 18 was linked to the downstream of the Tac promoter in the recombinant DNA pKGFMH2.

Ampicillin was added to L-broth (pH 7.2), which had been sterilized by autoclaving, to give a concentration of 50 μg/ml, cooled to 37°C, and inoculated with the transformant KGFMH2, followed by the culture at 37°C for 18 hours. Eighteen liters of a fresh preparation of the same culture medium was placed in a 20-ℓ jar fermenter, similarly sterilized as above, admixed with ampicillin, cooled to 37°C, and inoculated with one v/v % of the seed culture obtained in the above, followed by the culture at 37°C for 8 hours under aeration-agitation conditions. The resulting culture was centrifuged to collect the cultured cells which were then suspended in a mixture solution (pH 7.3) containing 150 mM sodium chloride, 16 mM disodium hydrogenphosphate, and 4 mM sodium dihydrogenphosphate, disrupted by ultrasonication, and centrifuged to remove cell disruptant, and this yielded an about two liters of a supernatant.

To an about two liters of the supernatant was added 10 mM phosphate buffer (pH 7.3) containing ammonium sulfate to give a 40% ammonium saturation. The resulting sediment was removed by centrifugation, and the supernatant

was mixed with ammonium sulfate to give an 85% ammonium saturation, allowed to stand at 4°C for 18 hours, and centrifuged at about 8,000 rpm for 30 min to obtain a newly formed sediment. The sediment thus obtained was dissolved in 10 mM phosphate buffer (pH 6.6) containing 1.5 M ammonium sulfate to give a total volume of about 1.300 ml, and the solution was filtered, and fed to a column packed with about 800 ml of "PHENYL SEPHAROSE CL-68", a gel commercialized by Pharmacia LKB Biotechnology AB, Uppsala, Sweden, followed by washing the column with a fresh preparation of the same buffer and feeding to the column a linear gradient buffer of ammonium sulfate decreasing from 1.5 M to O M in 10 mM phosphate buffer (pH 6.6) at an SV (space velocity) 1.5. Fractions eluted at around 1 M ammonium sulfate were pooled, concentrated using a membrane filter, and dialyzed against 10 mM phosphate buffer (pH 6.5) at 4°C for 18 hours. The dialyzed solution was fed to a column packed with about 55 ml of "DEAE-5PW", a gel commercialized by Pharmacia LKB Biotechnology AB, Uppsala, Sweden, which had been equilibrated with 10 mM phosphate buffer (pH 6.5). The column was washed with a fresh preparation of the same buffer, and fed with a linear gradient buffer of sodium chloride increasing from 0 M to 0.5 M in 10 mM phosphate buffer (pH 6.5) at SV 5.5, followed by collecting fractions eluted at around 0.2 M sodium chloride. Thereafter, the fractions were pooled and concentrated similarly as above up to give an about nine milliliters, followed by dialyzing the concentrate against PBS (phosphate buffered saline) at 4°C for 18 hours, and feeding the dialyzed solution to a column packed with "SUPERDEX 75", a gel commercialized by Pharmacia LKB Biotechnology AB, Uppsala, Sweden, which had been equilibrated with a fresh preparation of the same PBS. The column was fed with a fresh preparation of the same PBS to collect fractions with an IFN- γ inducing activity, and the fractions were pooled and concentrated with a membrane filter to obtain a purified mouse IL-18 in a yield of about 350 μg/ℓ culture.

According to the method in Japanese Patent Kokai No. 27,189/96, the purified mouse IL-18 was analyzed and determined for biological activity and physicochemical property as indicated below: Culturing mouse spleen cells, collected by a conventional manner, under different concentrations of the mouse IL-18 resulted in an IFN-γ production depending on the concentrations of the mouse IL-18, and this revealed that the mouse IL-18 has an activity of inducing IFN-γ production by spleen cells as an immunocompetent cell. In accordance with the method as reported by U. K. Laemmli in *Nature*, Vol. 227, pp. 680-685 (1970), the purified human IL-18 was subjected to SDS-PAGE under non-reducing conditions, resulting in a major band with an IFN-γ inducing activity at a position corresponding to 19,000±5,000 daltons. The N-terminal region of the mouse IL-18 contained the amino acid sequence of SEQ ID NO: 19 which corresponded to the N-terminal region of SEQ ID NO: 18.

With reference to Experiment 7, the biological activity of the IL-18 according to the present invention will be described in more detail, and Experiment 8 describes the cytotoxicity of the IL-18:

Experiment 7

10

15

20

25

30

35

40

45

50

Biological activity

Experiment 7-1

Induction of GM-CSF production

Using a heparinized syringe, blood was collected from a healthy volunteer and diluted two fold with serum-free RPMI 1640 medium (pH 7.4). The diluent was overlaid on a ficoll and centrifuged, and the collected lymphocytes were washed with RPMI 1640 medium (pH 7.4) supplemented with 10 v/v % fetal calf serum, and suspended in a fresh preparation of the same medium to give a cell density of 1 x 10⁶ cells/mI, followed by distributing the cell suspension to a 12-well microplate by two ml/well.

Using RPMI 1640 medium (pH 7.4) supplemented with 10 v/v % fetal calf serum, an IL-18 preparation obtained by the method in Experiment 1 was prepared into a one μ g/ml solution which was then distributed to the above microplate by 20-200 μ l/well. To the microplate was further added a fresh preparation of the same buffer, supplemented with 500 μ l/ml of Concanavalin A, by 10 μ l/well, followed by the incubation at 37°C for 48 hours in a 5 v/v % CO₂ incubator. After completion of the culture, supernatants in each well were sampled by 0.1 ml/well, and determined for GM-CSF content using a conventional enzyme immunoassay. In parallel, a culture system free of IL-18 as a control was provided and treated similarly as above. The data is in Table 1:

Table 1

	Table 1	
IL-18* (nM)	GM-CSF yield (pg/ml)	
0	510	
0.7	2,150	

Table 1 (continued)

IL-18* (nM)	GM-CSF yield (pg/ml)
2.8	3.050
5 6	3.950
Note: The symbol "*" means that IL-18 was navalin A.	s added to the culture system in the presence of 2.5 μg/ml of

The results in Table 1 indicate that lymphocytes as an immunocompetent cell produced GM-CSF depending on the concentration of IL-18 when contacted with IL-18 in the presence of Concanavalin A as a cofactor. It was also confirmed that all of the IL-18 preparations and functional equivalents thereof, which were obtained by the methods in Experiments 2 to 5, induced GM-CSF production even when used alone similarly as above. An IL-18 preparation obtained by the method in Experiment 6 was tested in accordance with Experiment 7-1 except that the human lymphocytes used in the experiment were replaced with spleen cells prepared from mouse by a conventional manner, revealing that the IL-18 preparation also induced GM-CSF production.

Experiment 7-2

5

10

15

20

25

30

35

40

45

55

Inhibition of osteoclast formation

Experiment 7-2(a)

As reported by T. J. Martin and K. W. Ng *in Journal* of Cellular *Biochemistry*, Vol. 56, pp. 357-366 (1994), it is considered requisite for contacting osteoclastic precursor cells, derived from hematopoietic stem cells, with osteoblasts or bone marrow stromas to generally differentiate osteoclastic precursor cells into mature osteoclasts. As described by G. D. Roodman in Endocrine *Reviews*, Vol. 17, No. 4, pp. 308-332 (1996), it is generally recognized that osteoclasts have characters of multinucleated cells, tartaric acid-resistant acid phosphatase (hereinafter abbreviated as "TRAP") activity, and a calcitonin receptor. In a co-culture system of osteoblasts and bone marrow cells as reported by Nobuyuki UDAGAWA et al., in *Journal of Experimental Medicine*, Vol. 182, pp. 1,461-1,468 (1995), these cells respond to factors such as 1α,25-dihydroxyvitamin D₃, prostaglandin E₂, adrenocortical hormone, interleukin 1, interleukin 6, and interleukin 11, to form osteoclast-like cells (hereinafter may be abbreviated as "OCL"). The formed OCL has characters of osteoclasts *in vivo*. Therefore, the co-culture system well reflects *in vitro* the processes of osteoclast formation in *vivo*. Using this system, experiments for osteoclast formation and osteoclastgenic inhibitory agents can be carried out.

The osteoclastgenic inhibitory activity of the IL-18 according to the present invention was studied using the above co-culture system. The osteoblasts used in this experiment were prepared in a conventional manner by treating a newborn mouse calvaria with 0.1 w/v % collagenase commercialized by Worthington Biochemical Co., Freefold, Australia, and 0.2 w/v % dispase commercialized by Godo Shusei Co., Ltd., Tokyo, Japan. The bone marrow cells were prepared from a mature mouse in a conventional manner. As a negative control, 2 x 10⁴ cells of a primary cell culture of osteoblasts and 5 x 105 cells of bone marrow cells were co-cultured in each well of a 48-well microplate containing 0.4 ml/well of α-MEM medium supplemented with 10 v/v % fetal calf serum (hereinafter designated as "Medium" throughout Experiment 4-2) at 37°C for seven days in a 5 v/v % CO₂ incubator. As a positive control, the above twotypes of cells were co-cultured similarly as in the negative control except that they were cultured in other wells containing 10⁻⁸M of 1α ,25-dihydroxyvitamin D₃ commercialized by Wako Pure Chemicals, Tokyo, Japan, and 10^{-7} M of prostaglandin E2 commercialized by Sigma Chemical Company, Missouri, USA. The aforesaid two-types of cells were cocultured similarly as in the positive control except that they were cultured in other wells containing 1a,25-dihydroxyvitamin D_3 commercialized by Wako Pure Chemicals, Tokyo, Japan, and prostaglandin E_2 commercialized by Sigma Chemical Company, Missouri, USA., in the same concentrations as used in the positive control, and a concentration of 0.01-10 ng/ml of an IL-18 preparation prepared by the method in Experiment 6. In every co-culture system, the media in each well were replaced with fresh preparations of the same media used in the co-culture systems on the 3rd day after the initiation of each culture. According to the method by Nobuyuki UDAGAWA in Journal of Experimental Medicine, Vol. 182, pp. 1,461-1,468 (1995), the cells on the 6th day after the initiation of each culture were fixed and stained based on TRAP activity, followed by counting the stained cells (hereinafter called "TRAP-positive cells") per well. Throughout Experiment 4-2, quadruplet wells under the same conditions were provided for each co-culture system, and the mean value for the TRAP-positive cells per well in each system was calculated. The results are in Table 2:

As shown in Table 2, the formation of TRAP-positive cells was not substantially observed in the negative control, but the distinct formation was observed in the positive control. In the co-culture systems, i.e., the positive control supplemented additionally with IL-18, the formation of TRAP-positive cells was inhibited depending on the concentration of IL-18, and the maximum inhibition, i.e., a level equal to that in the negative control, was found at eight ng/ml or more of IL-18. These data strongly indicates that IL-18 has a concrete activity of inhibiting OCL formation in *vitro* and also inhibits osteoclast formation.

Experiment 7-2(b)

As described hereinbefore, it was confirmed that there exist factors that induce the formation of osteoclast-like cells in the co-culture systems used throughout Experiment 7-2. Therefore, in this Experiment 7-2(b), it was studied whether the inhibitory activity of IL-18 on osteoclast formation observed in Experiment 7-2(a) was specific to some factors or not: the osteoclast-like cells were cultured by the same method as used in the negative control in Experiment 7-2(a) except for using a medium supplemented with 10^{-8} M $1\alpha.25$ -dihydroxyvitamin D_3 , 10^{-7} M prostaglandin E_2 , 200 ng/ml parathyroid hormone, 100 ng/ml interleukin 1, or 20 ng/ml interleukin 11. These culture systems were for positive controls. In parallel, the cells were cultured in other wells by the same method used in the positive controls except for using a medium containing 10 ng/ml of an IL-18 preparation obtained by the method in Experiment 6, in addition to any one of the above factors at the same concentration. After completion of the cultures. TRAP-positive cells in each well were counted, and the numbers were compared similarly as in Experiment 7-2(a). The results are in Table 3:

Table 3

5

10

15

20

25

30

35

40

45

50

55

Number of TRAP-positive cells per well*3	94	3	77	3	63	R	84	m	71	3
IL-18*2	1	+	ı	+		+	1	+	ľ	+
Osteoclast formation factor'l (concentration)	D_3 (10 ⁻⁴ M)		PGE, (10-7M)		PTH (200 ng/ml)		IL-11 (100 ng/ml)		IL-1 (20 ng/ml)	

prostaglandin E_2 , parathyroid hormone, interleukin-11, and interleukin-1 which were added to wells to give the concentrations as indicated in parentheses. The symbol "+" means that IL-18 was added to a well to give a concentration of 10 ng/ml, and the symbol "-" means that IL-18 was not added to. D_{3} , PGE,, PTH, IL-11, and IL-1 are respectively $1\alpha,25\text{-dihydroxyvitamin }D_{3}$, .. " *2: Note:

It shows a mean value of the data from quadruplet wells cultured under the same .; 3;

As shown in Table 3, a distinct formation of TRAP-positive cells was observed in every positive control, but the formation was almost completely inhibited in the presence of IL-18. This strongly indicates that IL-18 has a wide and general activity of inhibiting osteoclast formation independently of osteoclast-formation-related factors.

Experiment 7-2(c)

It was studied whether the osteoclastgenic inhibition by IL-18, confirmed in Experiments 7-2(a) and 7-2(b), was caused by the action of the IL-18-induced GM-CSF. For positive and negative controls, the same co-culture systems employed in Experiment 7-2(a) were used. Using other wells, the co-culture of osteoblasts and bone marrow cells was carried out similarly as the method used for the positive controls except for using a medium supplemented with 1α . 25-dihydroxyvitamin D_3 and prostaglandin E_2 at the same concentrations used in the positive control, and with (i) 10 μ g/ml of an anti-mouse GM-CSF polyclonal antibody commercialized by R&D Systems, Minnesota, USA, (ii) 10 μ g/ml of an IL-18 preparation obtained by the method in Experiment 6, (iii) (ii) plus 10 μ g/ml of an anti-mouse GM-CSF commercialized by R&D Systems, Minnesota, USA, or (v) (iv) plus 10μ g/ml of an anti-mouse GM-CSF polyclonal antibody. After completion of the culture, TRAP-positive cells in each well were counted, and the numbers were compared similarly as in Experiment 7-2(a). The data is shown in Table 4 where the symbols "i" to "v" coincide with those used in the co-culture systems other than the control systems.

Table 4

5

10

15

20

25

30

35

40

45

50

System*1 dateoclastgenic lib-18 3 GM-CSF 4 Anti-GM-CSF Number of TRAP-positive antibody*5 Anti-GM-CSF 4 Anti-GM-CSF vell*6 N - - - 3 P + - - 3 i + - - 3 ii + + - - 3 ix + - + 4 - 4 iv + - + - 4 - 4 v + - + + + - 4	Culture					
	em.1	Usteociastgenic factor ² 2	11-18 3	GM-CSF-4	Anti-GM-CSF antibody [*] 5	Number of TRAP-positive cells per well'6
+ + + + + + +		1	1	t	1	3
+ + + + + + + + + + + + + + + + + + + +		+	T.	l l	ı	122
+ + + + + + + + + + + + + + + + + + + +	Ą.	+	1	1	+	112
+ + + + + + + + + + + + + + + + + + + +	7	+	+			3
+ + +	11	+	+	ı	+	111
+	iv	+	ı	+		Ъ
		+	ı	+	+	106

controls, respectively, and the symbols "i" to "v" correspond "1; where the symbols "N" and "P" mean negative and positive Note:

to those in the five types co-culture systems used. where the symbol "+" means that $1\alpha,25-{\rm dihydroxyvitamin\ D_3}$ and prostaglandin E_2 were respectively added to a well to give respective concentrations of $10^{-8}M$ and $10^{-7}M$, and the symbol "-" means that these compounds were not added to. 2;

concentration of 10 ng/ml, and the symbol "-" means that IL-18 The symbol "+" means that IL-18 was added to a well to give a was not added to. 3;

concentration of 0.1 ng/ml, and the symbol "-" means that GM-CSF The symbol "+" means that GM-CSF was added to a well to give a 4;

was not added to. The symbol "+" means that an anti-GM-CSF polyclonal antibody was added to a well to give a concentration of $10~\mu g/ml$, and the symbol "-" means that the polyclonal antibody was not added to. ,2;

As shown in Table 4, the formation of TRAP-positive cells was almost completely inhibited by IL-18, cf., the co-culture system (ii), but the inhibition was almost completely inhibited by the addition of the anti-mouse polyclonal antibody, cf., the co-culture system (iii). Mouse GM-CSF exhibited an activity of inhibiting the formation of TRAP-positive cells similar to IL-18, cf., the co-culture system (iv), and the inhibition was almost completely inhibited by the addition of the anti-mouse GM-CSF polyclonal antibody, cf., the co-culture system (v). The sole use of the anti-mouse GM-CSF polyclonal antibody gave no influence on the formation of TRAP-positive cells, cf., the co-culture system (i). These data strongly indicates that the osteoclastgenic inhibition by IL-18 was due to the action of the IL-18-induced GM-CSF.

Experiment 8

10

20

25

30

35

40

45

55

Acute toxicity test

Eight-week-old mice were in a conventional manner injected percutaneously, or ally, or intraperitoneally with either of IL-18 preparations obtained by the methods in Experiments 1 to 6. The results showed that these IL-18 preparations had an LD₅₀ of about one mg/kg or more in mice independent of the route of administration. The data evidences that IL-18 can be incorporated into pharmaceuticals for warm-blooded animals in general and including humans without causing no serious side effects.

As described in *Nikkei Biotechnology Annual Report 1996*, pp. 498-499 (1995), published by Nikkei BP Publisher, Tokyo, Japan (1995), the IL-18-induced GM-CSF has not yet been clinically used in Japan, but applied clinically in USA and Europe. The fact would show that IL-18 has substantially no serious side effects. These facts indicate that the osteoclastgenic inhibitory agent according to the present invention can be successively administered to warm-blooded animals in general and including humans to induce osteoclast formation and exert a satisfactory therapeutic and/or prophylactic effect on osteoclast-related diseases without causing serious side effects.

The following Examples describe the present osteoclastgenic inhibitory agent according to the present invention:

Example 1

Liquid

Either of IL-18 preparations, obtained by the methods in Experiments 1 to 6, was dissolved in physiological saline containing one w/v % human serum albumin as a stabilizer to give a concentration of two mg/ml of the IL-18 preparation. The resulting solutions were in a conventional manner membrane filtered for sterilization into liquids.

The liquids have a satisfactory stability and can be arbitrarily used as ingredients for cell culture and agents in the form of an injection, ophthalmic solution, or collunarium for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Example 2

Dry agent

Fifty milligrams of either of IL-18 preparations, obtained by the methods in Experiments 1 to 6, was dissolved in 100 ml of physiological saline containing one w/v % purified gelatin as a stabilizer. The solutions thus obtained were in a conventional manner membrane filtered for sterilization, distributed to vials by one milliliter, lyophilized, and sealed with caps

The products have a satisfactory stability and can be arbitrarily used as ingredients for cell culture and agents in the form of a dry injection for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Example 3

Dry agent

Fifty milligrams of either of IL-18 preparations, obtained by the methods in Experiments 1 to 6, was dissolved in 100 ml of physiological saline containing one w/v % trehalose as a stabilizer. The solutions were in a conventional manner membrane filtered for sterilization, distributed to vials by one milliliter, lyophilized, and sealed with caps.

The products have a satisfactory stability and can be arbitrarily used as ingredients for cell culture and agents in the form of a dry injection for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Example 4

Ointment

5

10

15

25

30

35

40

45

50

"HIVIS WAKO GEL® 104", a carboxyvinylpolymer commercialized by Wako Pure Chemical Industries, Ltd., Tokyo, Japan, and a high-purity trehalose were dissolved in a sterilized distilled water to give respective concentrations of 1.4 w/w % and 2.0 w/w %, and the solution was mixed to homogeneity with either of IL-18 preparations obtained by the methods in Experiments 1 to 6, and adjusted to pH 7.2 to obtain a paste containing about one mg of an IL-18 preparation per g of the product.

Each product thus obtained has a satisfactory spreadability and stability and can be arbitrarily used as an agent in the form of an ointment for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Example 5

Tablet

"FINETOSE®*, an anhydrous crystalline α-maltose powder commercialized by Hayashibara Biochemical Laboratories, Inc., Okayama, Japan, was mixed to homogeneity with either of IL-18 preparations, obtained by the methods in Experiments 1 to 6, and "LUMIN" or 1-1'-1*-trihepthyl-11-chinolyl(4)*4*4'-penthamethinchynocyanine-1,1"-dijodide. The mixtures were in a conventional manner tabletted to obtain tablets, about 200 mg weight each, containing an about two milligrams of either of the IL-18 preparations and an about two milligrams of LUMIN per tablet.

The products have a satisfactory swallowability, stability, and cell-activating activity and can be arbitrarily used as agents in the form of a tablet for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

As described above, the osteoclastgenic inhibitory agent according to the present invention effectively inhibits osteoclast formation. Therefore, the agent can be arbitrarily used as an ingredient for cell culture and agents for regulating bone resorption and for osteoclast-related diseases, directed to treat and/or prevent hypercalcemia, osteoclastoma, osteoporosis, etc.

Thus the present invention with these useful activities and functions is a significant invention that would greatly contribute to this field.

While there has been described what is at present considered to be the preferred embodiments of the invention, it will be understood the various modifications may be made therein, and it is intended to cover in the appended claims all such modifications as fall within the true spirits and scope of the invention.

Annex to the description

5	SEQUENCE LISTING
	(1) INFORMATION FOR SEQ ID NO: 1:
0	<pre>(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 6 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear</pre>
	(ii)MOLECULE TYPE: peptide
5	(v)FRAGMENT TYPE: internal fragment
	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 1:
20	Asn Asp Gln Val Leu Phe 1 5
25	<pre>(2) INFORMATION FOR SEQ ID NO: 2: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 6 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear</pre>
	(ii)MOLECULE TYPE: internal fragment
30	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 2:
	Phe Glu Asp Met Thr Asp 1 5
	(3) INFORMATION FOR SEQ ID NO: 3:
35	(i)SEQUENCE CHARACTERISTICS:(A)LENGTH: 7 amino acids(B)TYPE: amino acid(D)TOPOLOGY: linear
40	(ii)MOLECULE TYPE: peptide
	(v)FRAGMENT TYPE: internal fragment
45	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 3:
45	Phe Lys Leu Ile Leu Lys Lys
	(4) INFORMATION FOR SEQ ID NO: 4:
50	(i)SEQUENCE CHARACTERISTICS:(A)LENGTH: 5 amino acids(B)TYPE: amino acid(D)TOPOLOGY: linear
55	(ii)MOLECULE TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 4: 5 Met Tyr Lys Asp Ser (5) INFORMATION FOR SEQ ID NO: 5: (i)SEQUENCE CHARACTERISTICS: 10 (A)LENGTH: 5 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear (ii) MOLECULE TYPE: internal fragment 15 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 5: Ser Thr Leu Ser Cys 5 20 INFORMATION FOR SEQ ID NO: 6: (6) (i)SEQUENCE CHARACTERISTICS: 25 (A)LENGTH: 157 amino acids (B) TYPE: amino acid (D)TOPOLOGY: linear (ii) MOLECULE TYPE: peptide 30 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 6: Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 35 20 25 Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 60 Ser Val Lys Cys Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 40 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 45 105 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 120 125 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 145 150 (7) INFORMATION FOR SEQ ID NO: 7: (i) SEQUENCE CHARACTERISTICS: (A)LENGTH: 157 amino acids

50

	(B)TYPE: amino acid (D)TOPOLOGY: linear																
5		(i:	L)MOI	LECUI	LE T	PE:	pep	tide									
		(x:	L)SE(QUEN	CE DE	ESCR:	IPTIC	ON: S	SEQ :	ID NO	0: 7	:					
10	Asn 1	Phe	Gly	Arg	Leu 5	His	Cys	Thr	Thr	Ala 10	Val	Ile	Arg	Asn	Ile 15	Asn	
		Gln		20					25					30	_		
		Asp	35					40					45				
15		Met 50					55					60					
	65	Lys				70					75					80	
20		Phe			85					90			_		95		
		Leu Glu		100					105					110			
		Asp	115					120					125		_		
25		130					135					140		ASII	GIY	wah	
	Lys Ser Val Met Phe Thr Leu Thr Asn Leu His Gln Ser 145 150 155																
	(8)INFORMATION FOR SEQ ID NO: 8:																
30	(1)SEQUENCE CHARACTERISTICS: (A)LENGTH: 471 base pairs																
	(B)TYPE: nucleic acid (C)STRANDEDNESS: double																
35	(C)STRANDEDNESS: double (D)TOPOLOGY: linear																
		•			E TY			1									
		(V 1	(A	ORG	L SC ANIS L TY	M: h	numar										
40		(f x)FEA			FE.	TTAG	3.L									
		(2.2	(A	AAN (E/KE				Lđe								
45					NTIF				HOD:	Ε							
		(xi)SEQ	UENC	E DE	SCRI	PTIC)N: S	SEQ I	D NO): 8:	:					
	TAC TT Tyr Ph			s Le					u Se	r Va				n Le	u As		48
50	GAC CA									T CG					A GA		96
	Asp Gl		20				-	25			-		30			~	, , ,
	ATG AC	r GA	I IC	ı GAI	U TG	ı AG	n GA	r AA	r GC	A CC	د دن	G AC	C AT	A TT	г АТ	T.	144

	Met	Thr	35	Ser	Asp	Cys	Arg	Asp 40	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile	
5		50	ATG Met	. 4	-10		55	GIII	FIO	Arg	GTĀ	Met	Ala	Val	Thr	Ile	192
	65		AAG Lys	-10		70	ATT Ile	Ser	LIII	Leu	Ser	Cys	Glu	Asn	Lys	Ile	240
10			TTT Phe		85		NSII	FLO	PLO	ASP	AAC Asn	Ile	Lys	Asp	Thr	Lys	288
1.5		-	ATC Ile	100			0111	AL 9	105	GTC Val	Pro	GIY	His	Asp	Asn	Lys	336
15			TTT Phe 115				Jer	120	GAA Glu	GTA	Tyr	Phe	Leu	Ala	Cys	Glu	384
20	_	130	AGA Arg				135	CTC Leu	TIE	reu	Lys	Lys	Glu	GAT Asp	GAA Glu	TTG Leu	432
	GGG Gly 145	GAT Asp	AGA Arg	TCT Ser	ATA	ATG Met 150	TTC Phe	ACT Thr	GTT Val	CAA Gln	AAC Asn 155		GAC Asp				471
25	(9)		ORMA														
	(i)SEQUENCE CHARACTERISTICS:(A)LENGTH: 11 amino acids(B)TYPE: amino acid(D)TOPOLOGY: linear																
30	(ii)MOLECULE TYPE: peptide																
		(v)	FRAGI	MENT	TYP	E: N-	-ter	mina	l fra	agmer	nt					-	
35)SEQ														
	Met 1	Tyr I	Phe (Gly I	iys 1	Leu (Glu S	Ser I		Leu S LO	Ser						
40	(10)	INFO	ORMAT	CION	FOR	SEQ	ID N	10: 1	.0:								
		(i)S	(B)	LENG	TH: : an	ACTE 10 a ino : li	mino acid	aci	ds								
45		(ii)	MOLE	CULE	TYP	E: p	epti	de									
		(v)F	RAGM	ENT	TYPE	: C-	term	inal	fra	gmen	t						
50		(xi)	SEQU	ENCE	DES	CRIP	TION	: SE	Q ID	NO:	10:						
	Ser I l	le M	et P	he T	hr V	al G	ln A	sn G	lu A								

	(11) INFORMATION FOR SEQ ID NO: 11:
5	<pre>(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 13 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear</pre>
	(ii)MOLECULE TYPE: peptide
10	(v)FRAGMENT TYPE: N-terminal fragment
	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 11:
15	Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg 1 5 10
	(12) INFORMATION FOR SEQ ID NO: 12:
20	<pre>(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 14 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear</pre>
	(ii)MOLECULE TYPE: peptide
25	(v)FRAGMENT TYPE: internal fragment
	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 12:
30	Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg 1 10
	(13) INFORMATION FOR SEQ ID NO: 13:
35	<pre>(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 17 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear</pre>
	(ii)MOLECULE TYPE: peptide
40	(v)FRAGMENT TYPE: internal fragment
	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 13:
45	Ile Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 1 10 15
	(14)INFORMATION FOR SEQ ID NO: 14:
50	<pre>(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 471 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: double (D)TOPOLOGY: linear</pre>

		(i:	L)MOI	LECUI	LE TY	PE:	CDNA	4									
5		(1)	(E	A)NAN B)LOC	ME/KE	EY: m DN: 1 FICAT	47	71		S							
		(xi	.)SE(): 14	:					
10	71. C	mmm	000		O m m	<i>-</i>	m ~ m										
	Tyr 1	Phe	GGC Gly	Lys	Leu 5	Glu	Ser	ГЛЗ	Leu	Ser 10	Val	Ile	Arg	Asn	Leu 15	Asn	48
15	GAC Asp	CAA Gln	GTT Val	CTC Leu 20	TTC Phe	ATT Ile	GAC Asp	CAA Gln	GGA Gly 25	AAT Asn	CGG Arg	CCT Pro	CTA Leu	TTT Phe 30	GAA Glu	GAT Asp	9 6
	ATG Met	ACT Thr	GAT Asp	TCT	GAC Asp	TCT Ser	AGA Arg	Asp	AAT	GCA Ala	CCC Pro	CGG Arg	Thr	ATA	TTT Phe	ATT Ile	144
20	ATA Ile	AGT Ser	35 ATG Met	TAT Tyr	AAA Lys	GAT Asp	AGC Ser	40 CAG Gln	CCT Pro	AGA Arg	GGT Glv	ATG Met	45 GCT Ala	GTA Val	ACT	ATC	192
	TCT	50 GTG	AAG	TCT	GAG	AAA	55 ATT	TCA	ACT	CTC	TCC	60 GCT	GAG	AAC	AAA	АТТ	240
	65		Lys TTT			70					75					80	200
?5	Ile	Ser	Phe	Lys	Glu 85	Met	Asn	Pro	Pro	Asp 90	Asn	Ile	Lys	Asp	Thr 95	Lys	288
	AGT Ser	GAC Asp	ATC Ile	ATA Ile 100	TTC Phe	TTT Phe	CAG Gln	AGA Arg	AGT Ser 105	GTC Val	CCA Pro	GGA Gly	CAT His	GAT Asp 110	AAT Asn	AAG Lys	336
30	ATG Met	CAA Gln	TTT	GAA	TCT Ser	TCA Ser	TCA Ser	Tyr	GAA	GGA Gly	TAC Tyr	TTT Phe	Leu	GCT	TGT Cys	GAA Glu	384
	AAA Lys	GAG Glu	115 AGA Arg	GAC Asp	CTT Leu	TTT Phe	AAA Lys	120 CTC Leu	ATT Ile	TTG Leu	AAA Lys	AAA Lys	125 GAG Glu	GAT Asp	GAA Glu	TTG Leu	432
15	GGG	130 GAT	AGA Arg	TCT	ATA	ATG	135 TTC	ACT	GTT	CAA	AAC	140 GAA	GAC				471
	145	vəħ	Arg	Ser	116	150	FILE	1111	AGI	GIN	155	GIU	Asp				
0	(15)		FORM														
		(1	(E	A)LEN	GTH:	: 10 amino	amir aci	no ac id									
5		(ii	L)MOI	LECUI	E TY	PE:	pept	ide									
		(v))FRAC	MENT	TYF	E: N	1-tei	cmina	al fr	agme	ent						
0)SEC): 15	i :					
	Tyr 1	Phe	Gly	Lys	Leu 5	Glu	Ser	Lys	Leu	Ser 10							

(16) INFORMATION FOR SEQ ID NO: 16:

5		(i)	(E	A)LEN B)TYE C)STE	IGTH: PE: r RANDE	RACT 471 nucle EDNES SY: 1	bas ic a SS: d	se pa acid loubl	irs								
10		(ii	L)MOI	LECUI	E TY	PE:	CDNA	A									
		(i)	(E	A) NAN B) LOC	1E/KE	EY: n ON: 1	4.	71		s							
15		(x:	i)SE(QUENC	CE DE	ESCRI	PTIC)N: S	SEQ I	מ ס): 16	5:					
	TAC Tyr 1	TTT Phe	GGC Gly	AAG Lys	CTT Leu 5	GAA Glu	TCT Ser	AAA Lys	TTA Leu	TCA Ser 10	GTC Val	ATA Ile	AGA Arg	AAT Asn	TTG Leu 15	AAT Asn	48
20	GAC Asp	CAA Gln	GTT Val	CTC Leu 20	TTC Phe	ATT Ile	GAC Asp	CAA Gln	GGA Gly 25	AAT Asn	CGG Arg	CCT Pro	CTA Leu	TTT Phe 30	GAA Glu	GAT Asp	96
			GAT Asp 35														144
25	ATA Ile	AGT Ser 50	ATG Met	TAT Tyr	AAA Lys	GAT Asp	AGC Ser 55	CAG Gln	CCT Pro	AGA Arg	GGT Gly	ATG Met 60	GCT Ala	GTA Val	ACT Thr	ATC Ile	192
30	TCT Ser 65	GTG	AAG Lys	TCT Ser	GAG Glu	AAA Lys 70	ATT Ile	TCA Ser	ACT Thr	CTC Leu	TCC Ser 75	GCT Ala	GAG Glu	AAC Asn	AAA Lys	ATT Ile 80	240
			TTT Phe														288
35			ATC Ile														336
	Met	Gln	TTT Phe 115	Glu	Ser	Ser	Ser	Tyr 120	Glu	Gly	Tyr	Phe	Leu 125	Ala	Ser	Glu	384
40	Lys	Glu 130	AGA Arg	Asp	Leu	Phe	Lys 135	Leu	Ile	Leu	Lys	Lys 140	Glu	GAT Asp	GAA Glu	TTG Leu	432
	GGG Gly 145	GAT Asp	AGA Arg	TCT Ser	ATA Ile	ATG Met 150	TTC Phe	ACT Thr	GTT Val	CAA Gln	AAC Asn 155	GAA Glu	GAC Asp				471
4 5	(17	•		UENCI A)LEI	E CHA		reri:	STIC: base	s:	rs							
50			(C)ST	RAND	EDNE:	ss:	doub	le								

	(ii)MOLECULE TYPE: genomic DNA	
5	<pre>(vi)ORIGINAL SOURCE: (A)ORGANISM: human (G)CELL TYPE: placenta</pre>	
10	<pre>(ix)FEATURE: (A)NAME/KEY: 5´ UTR (B)LOCATION: 13 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: leader peptide (B)LOCATION: 482</pre>	
15	(C)IDENTIFICATION METHOD: S (A)NAME/KEY: intron (B)LOCATION: 831453 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: leader peptide (B)LOCATION: 14541465	
20	(C)IDENTIFICATION METHOD: S (A)NAME/KEY: intron (B)LOCATION: 14664848 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: leader peptide	
25	(B)LOCATION: 48494865 (C)IDENTIFICATION METHOD: S (A)NAME/KEY: mat peptide (B)LOCATION: 48664983 (C)IDENTIFICATION METHOD: S	
30	(A)NAME/KEY: intron (B)LOCATION: 49846317 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: mat peptide (B)LOCATION: 63186451	
s <i>s</i>	(C)IDENTIFICATION METHOD: S (A)NAME/KEY: intron (B)LOCATION: 645211224 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: mat peptide (B)LOCATION: 1122511443	
0	(C)IDENTIFICATION METHOD: S (A)NAME/KEY: 3° UTR (B)LOCATION: 1144411464 (C)IDENTIFICATION METHOD: E	
	(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 17:	
5	AAG ATG GCT GCT GAA CCA GTA GAA GAC AAT TGC ATC AAC TTT GTG GCA Met Ala Ala Glu Pro Val Glu Asp Asn Cys Ile Asn Phe Val Ala -35 -30 -25	48
	ATG AAA TTT ATT GAC AAT ACG CTT TAC TTT ATA G GTAAGG CTAATGCCAT Met Lys Phe Ile Asp Asn Thr Leu Tyr Phe Ile Ala -20 -15 -10	98
0	AGAACAAATA CCAGGTTCAG ATAAATCTAT TCAATTAGAA AAGATGTTGT GAGGTGAACT ATTAAGTGAC TCTTTGTGTC ACCAAATTTC ACTGTAATAT TAATGGCTCT TAAAAAAATA GTGGACCTCT AGAAATTAAC CACAACATGT CCAAGGTCTC AGCACCTTGT CACACCACGT GTCCTGGCAC TTTAATCAGC AGTAGCTCAC TCTCCAGTTG GCAGTAAGTG CACATCATGA	158 218 278 338

	AAATCCCAGT	TTTCATGGGA	AAATCCCAGT	TTTCATTGGA	TTTCCATGGG .	AAAAATCCCA	398
	GTACAAAACT	GGGTGCATTC	AGGAAATACA	ATTTCCCAAA	GCAAATTGGC .	AAATTATGTA	458
					TTATGTAAAT		518
5					CTGGAGTGCA		578
					TCTCCTGCCT		638
					ATTTTTGGGT		698
					TCCTGATCTC		758
					ACCACCACAC		818
	A A MMC A MMC M	TATCATTA AT	CTCCTCTCAA	CAATTTCCCT	TCATTTGAAA	CTTTCCCTTC	878
10					ACCCCATCTC		938
					CTGAAGCAGG		998
	TGCAAAATAT	CCTGTGGACA	CCICCIACCI	AMCAMCCCAC	CCCTACACTC	CACCECCACII	
	GAGCCTAGGA	ATTTGAGCCT	GCAGTGAGCT	ATGATCCCAC	CCCTACACTC	AMM ACMMCAC	1058
					ATAAAAAATT		1118
	TTTTCTTAGG	TGACTTTCCG	TTTAAGCAAT	AAATTTAAAA	GTAAAATCTC	TAATTTTAGA	1178
15	AAATTTATT	TTAGTTACAT	ATTGAAATTT	TTAAACCCTA	GGTTTAAGTT	TTATGTCTAA	1238
					TGGGCCTTTT		1298
					AAAAATAGGA		1358
					TGAGTCGAAA		1418
	TACATATTCT	GTTTCTCTCT	TTTTCCCCCT		GAA GAT GAT		1470
20					Glu Asp Asp	o Glu	
20				-10			
	GTAGAAATGA	ATTTATTTT	CTTTGCAAAC	TAAGTATCTG	CTTGAGACAC	ATCTATCTCA	1530
	CCATTGTCAG	CTGAGGAAAA	AAAAAAATGG	TTCTCATGCT	ACCAATCTGC	CTTCAAAGAA	1590
	ATGTGGACTC	AGTAGCACAG	CTTTGGAATG	AAGATGATCA	TAAGAGATAC	AAAGAAGAAC	1650
	CTCTAGCAAA	AGATGCTTCT	CTATGCCTTA	AAAAATTCTC	CAGCTCTTAG	AATCTACAAA	1710
25	ATAGACTTTG	CCTGTTTCAT	TGGTCCTAAG	ATTAGCATGA	AGCCATGGAT	TCTGTTGTAG	1770
	GGGGAGCGTT	GCATAGGAAA	AAGGGATTGA	AGCATTAGAA	TTGTCCAAAA	TCAGTAACAC	1830
	CTCCTCTCAG	AAATGCTTTG	GGAAGAAGCC	TGGAAGGTTC	CGGGTTGGTG	GTGGGGTGGG	1890
	GCAGAAAATT	CTGGAAGTAG	AGGAGATAGG	AATGGGTGGG	GCAAGAAGAC	CACATTCAGA	1950
	GGCCAAAAGC	TGAAAGAAAC	CATGGCATTT	ATGATGAATT	CAGGGTAATT	CAGAATGGAA	2010
	GTAGAGTAGG	AGTAGGAGAC	TGGTGAGAGG	AGCTAGAGTG	ATAAACAGGG	TGTAGAGCAA	2070
30	GACGTTCTCT	CACCCAAGA	TGTGAAATTT	GGACTTTATC	TTGGAGATAA	TAGGGTTAAT	2130
	TAAGCACAAT	ATGTATTAGC	TAGGGTAAAG	ATTAGTTTGT	TGTAACAAAG	ACATCCAAAG	2190
	ATACAGTAGC	TGAATAAGAT	AGAGAATTTT	TCTCTCAAAG	AAAGTCTAAG	TAGGCAGCTC	2250
	AGAAGTAGTA	TGGCTGGAAG	CAACCTGATG	ATATTGGGAC	CCCCAACCTT	CTTCAGTCTT	2310
	GTACCCATCA	TCCCCTAGTT	GTTGATCTCA	CTCACATAGT	TGAAAATCAT	CATACTTCCT	2370
	GGGTTCATAT	CCCAGTTATC	AAGAAAGGGT	CAAGAGAAGT	CAGGCTCATT	CCTTTCAAAG	2430
35	ACTCTAATTG	GAAGTTAAAC	ACATCAATCC	CCCTCATATT	CCATTGACTA	GAATTTAATC	2490
	ACATGGCCAC	ACCAAGTGCA	AGGAAATCTG	GAAAATATAA	TCTTTATTCC	AGGTAGCCAT	2550
	ATGACTCTTT	AAAATTCAGA	AATAATATAT	TTTTAAAATA	TCATTCTGGC	TTTGGTATAA	2610
	AGAATTGATG	GTGTGGGGTG	AGGAGGCCAA	AATTAAGGGT	TGAGAGCCTA	TTATTTTAGT	2670
	TATTACAACA	AATGATGGTG	TCATGAATTA	AGGTAGACAT	AGGGGAGTGC	TGATGAGGAG	2730
40	CTGTGAATGG	ATTTTAGAAA	CACTTGAGAG	AATCAATAGG	ACATGATTTA	GGGTTGGATT	2790
40	TGGAAAGGAG	AAGAAAGTAG	AAAAGATGAT	GCCTACATTT	TTCACTTAGG	CAATTTGTAC	2850
	CATTCACTCA	AATAGGGAAC	ACAGGAGGAA	GAGCAGGTTT	TGGTGTATAC	AAAGAGGAGG	2910
	ATCCATGACG	CATTTCGTTT	TGGATCTGAG	ATGTCTGTGG	AACGTCCTAG	TGGAGATGTC	2970
	CACAAACTCT	TOTACATOTC	GTTCTGAGTT	CAGGACACAG	ATTTGGGCTG	GAGATAGAGA	3030
	TATTE ACCC	TTATACATAC	AAATGGCATT	TGAATCTATA	GAGATAAAA	GACACATCAG	3090
45	TATIGIAGG	TIMINOMING	CACCAAAACC	СААСТАСТСТ	GCTGGGGGGA	ATACCTACAT	3150
	MUSAAATGIG	CACMACAAAC	AACCEBAEAA	ACAACAGAGA	GCAGACTAAC	CAAAAGGGGA	3210
	CARCARAGGATG	CAGIAGAAAG	CCACCCACTC	CCACGAGAGA	ATTTCAAGAT	TGAGGGGATA	3270
	CCMCMMCMCM	TO A A TOTAL COLOR	ACCCTTCACACA	ATC A ACCCCC	AGAACACAGC	ΨΨΨΨΑΘΑΨΨΨ	3330
	GGTGTTGTGT	A COMMOCOMOS	MCMC3 CMC3 3	A CCA CCTTCA	TGGTGAAATG	GAGGCAGAGC	3390
	AGCAACAAGG	AGTITGGTGA	A CTCAGIGAA	ACCECANCAN	ATGATACAGA	TA A TOTOLOGICA	3450
50	CAGATTGCAA	TGAGTGAAAC	CONCONCINC	ANGIGANGAN	VIGWIYCVGV	CACATTOTICO	3510
	TAAAAGCTTG	GUTGTTAAAA	GGAGGAGAGA	CACCEMENCES	GCTGCAAAGT	MAMCA A A A MA	3570
	TGATGGAGCA	GTTTTAAATC	TUAAAATAAA	GAGCTTTGTG	CTTTTTTGAT	INIGNAMATA	3370

	ATGTGTTAAT TGTAACTAAT TGAGGCAATG AAAAAAGATA ATAATATGAA AGATAAAAAT	0.600
	ATAAAAACCA CCCAGAAATA ATGATAGCTA CCATTTTGAT ACAATATTC TACACTCCTT	3630
	TCTATGTATA TATACAGACA CAGAAATGCT TATATTTTTA TTAAAAGGGA TTGTACTATA	3690
5	CCTAAGCTGC TTTTCTAGT TAGTGATATA TATGGACATC TCTCCATGGC AACGAGTAAT	3750
	TGCAGTTATA TTAAGTTCAT GATATTTCAC AATAAGGGCA TATCTTTGCC CTTTTTATTT	3810
	ANCINTET TANTECTES ACCORDED CONCERN CO	3870
	AATCAATTCT TAATTGGTGA ATGTTTGTTT CCAGTTTGTT GTTGTTATTA ACAATGTTCC	3930
	CATAAGCATT CCTGTACACC AATGTTCACA CATTTGTCTG ATTTTTTCTT CAGGATAAAA	3990
	CCCAGGAGGT AGAATTGCTG GGTTGATAGA AGAGAAAGGA TGATTGCCAA ATTAAAGCTT	4050
10	CAGTAGAGGG TACATGCCGA GCACAAATGG GATCAGCCCT AGATACCAGA AATGGCACTT	4110
	TCTCATTTCC CCTTGGGACA AAAGGGAGAG AGGCAATAAC TGTGCTGCCA GAGTTAAATT	4170
	TGTACGTGGA GTAGCAGGAA ATCATTTGCT GAAAATGAAA ACAGAGATGA TGTTGTAGAC	4230
	GTCCTGAAGA GAGCAAAGAA AATTTGAAAT TGCGGCTATC AGCTATGGAA GAGAGTGCTC	4290
	AACTGGAAAA CAAAAGAAGT ATTGACAATT GGTATGCTTG TAATGGCACC GATTTGAACC	4350
	CTTGTGCCAT TGTTCACCAG CAGCACTCAG CAGCCAAGTT TGGAGTTTTG TAGCAGAAAC	4410
15	ACAAATAAGT TAGGGATTTA ATATCCTGGC CAAATGGTAG ACAAATGA CTCTGACATG	
	CAGCTGCACA GGGAAGGAAG GGAAGACGGG AAGAGGTTAG ATAGGAAATA CAAGAGTCAG	4470
	GAGACTGGAA GATGTTGTGA TATTTAAGAA CACATAGAGT TGGAGTAAAA GTGTAAGAAA	4530
	ACTAGAAGGG TAAGAGACCG GTCAGAAAGT AGGCTATTTG AAGTTAACAC TTCAGAGGCA	4590
	GAGTAGTTCT GAATGGTAAC AAGAAATTGA GTGTGCCTTT GAGAGTAGGT TAAAAAAACAA	4650
	TAGGCAACTT TATTGTAGCT ACTTCTGGAA CAGAAGATTG TCATTAATAG TTTTAGAAAA	4710
20	CTABLETA TACCAMACTE ACTICIONA CAGAGATTI TCATTAATAG TTTTAGAAAA	4770
	CTAAAATATA TAGCATACTT ATTTGTCAAT TAACAAAGAA ACTATGTATT TTTAAATGAG ATTTAATGTT TATTGTAG AA AAC CTG GAA TCA GAT TAC TTT GGC AAC CTT	4830
		4880
	Glu Asn Leu Glu Ser Asp Tyr Phe Gly Lys Leu	
	-5 1 5	
25	GAA TCT AAA TTA TCA GTC ATA AGA AAT TTG AAT GAC CAA GTT CTC TTC	4928
25	Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe	
	10 15 20	
	ATT GAC CAA GGA AAT CGG CCT CTA TTT GAA GAT ATG ACT GAT TCT GAC	4976
	Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp Met Thr Asp Ser Asp	
	25 30 35 35	
30	TGT AGA G GTATTTTT TTAATTCGCA AACATAGAAA TGACTAGCTA CTTCTTCCCA	5032
50	Cys Arg Asp	
	40	
	TTCTGTTTTA CTGCTTACAT TGTTCCGTGC TAGTCCCAAT CCTCAGATGA AAAGTCACAG	5092
	GAGTGACAAT AATTTCACTT ACAGGAAACT TTATAAGGCA TCCACGTTTT TTAGTTGGGG	5152
	TAAAAAATTG GATACAATAA GACATTGCTA GGGGTCATGC CTCTCTGAGC CTCCCTTTCA	5212
35	ATCACCAATC CCTTTATTGT GATTGCATTA ACTGTTTAAA ACCTCTATAG TTGGATGCTT	5272
	AATCCCTGCT TGTTACAGCT GAAAATGCTG ATAGTTTACC AGGTGTGGTG GCATCTATCT	5332
	GTAATCCTAG CTACTTGGGA GGCTCAAGCA GGAGGATTGC TTGAGGCCAG GACTTTGAGG	
	CTGTAGTACA CTGTGATCGT ACCTGTGAAT AGCCACTGCA CTCCAGCCTG GGTGATATAC	5392 5 4 52
	AGACCTTGTC TCTAAAATTA AAAAAAAAA AAAAAAAAAC CTTAGGAAAG GAAATTGATC	
	AAGTCTACTG TGCCTTCCAA AACATGAATT CCAAATATCA AAGTTAGGCT GAGTTGAAGC	5512
40	AGTGAATGTG CATTCTTTAA AAATACTGAA TACTTACCTT AACATATATT TTAAATATTT	5572
	TATTTAGCAT TTAAAAGTTA AAAACAATCT TTTAGAATTC ATATCTTTAA AATACTCAAA	5632
	ANGUEGO COUNTY CONTROL OF THE ANGUEGO CONTROL OF THE CONTROL OF TH	5692
	AAAGTTGCAG CGTGTGTTT GTAATACACA TTAAACTGTG GGGTTGTTTG TTTGTTTGAG	5752
	ATGCAGTTTC ACTCTGTCAC CCAGGCTGAA GTGCAGTGCA	5812
	CTCACTACAA CCTCCACCTC CCACGTTCAA GCGATTCTCA TGCCTCAGTC TCCCGAGTAG	5872
45	GTGGGATTAC AGGCATGCAC CACTTACACC CGGCTAATTT TTGTATTTTT AGTAGAGCTG	5932
	GGGTTTCACC ATGTTGGCCA GGCTGGTCTC AAACCCCTAA CCTCAAGTGA TCTGCCTGCC	5992
	TCAGCCTCCC AAACAAACAA ACAACCCCAC AGTTTAATAT GTGTTACAAC ACACATGCTG	6052
	CAACTTTTAT GAGTATTTTA ATGATATAGA TTATAAAAGG TTGTTTTTAA CTTTTAAATG	6112
	CTGGGATTAC AGGCATGAGC CACTGTGCCA GGCCTGAACT GTGTTTTTAA AAATGTCTGA	6172
	CCAGCTGTAC ATAGTCTCCT GCAGACTGGC CAAGTCTCAA AGTGGGAACA GGTGTATTAA	6232
50	GGACTATCCT TTGGTTAAAT TTCCGCAAAT GTTCCTGTGC AAGAATTCTT CTAACTAGAG	6292
	TTCTCATTTA TTATATTTAT TTCAG AT AAT GCA CCC CGG ACC ATA TTT ATT	6343
	Asp Asn Ala Pro Arg Thr Ile Phe Ile	0040
	one one of the file	

			40		45	
	አመን እርመ አመር (מאת אא כא		יכיד אכא מפיד	ATG GCT GTA ACT A	rc 6391
	Ala AGI AIG	TAI AAA OA	n Sor Cla	ero Ara Gly	Met Ala Val Thr I	10 0071
5	50	TAT TAP WE	55 551 5111 1	TO ALG GLY	60	.16
3	3U	TOT CAC AA		כיי כייר יירר	TGT GAG AAC AAA A'	rr 6439
					Cys Glu Asn Lys I	
	65	cys Gid Ly 70		75		30
					CAATCATGTT AATATA	
	Ile Ser Phe		nc i onocci i	ACTITOTITI	CIMICATOTT THITTE	111C/1 0430
10			ב ב תר⊃ידיתיים בייתי	מיימיימיימיים	AGTAATGTAA TTAGAA	AACT 6556
					ACAAGAAGCA GAGAAC	
					TTGTGATAAT GATGGT	
					TTATGACCTG CATCTC	
	CIGAGCCIGI C	MCMGGGGMA	TAATAATCTA	CAATACATAT	GTGAGTTATA CATTTA	AGAA 6796
	MANGAGIC I	TIMINUMAN	TCACTTCTCC	TATCALCALAT	GAAGCTAATT ATCCTT	CTAT 6856
15	AMERICAN CAC C	TTTCCAGAA	TATCATAATA	TATOAAGAAT	TAGTTGTTTT GTTGCT	GATC 6916
	CTTACCCTAA C	TOTALL	CAACCTTCAG	CTTCCAGTTG	ATGTATGTTA TTTTTA	ATGT 6976
	TARGUCIAA G	AATAAAACT	TATCACATCA	CCTCTAAAAG	TAATGCTATA ATTATC	TTCA 7036
	ACCACCATA A	AACTATTTC	TEGECTETAC	TTTTTCTCTA	TTATTCTCCA TTATTA	TTCT 7096
	CTATTATTATT	CTCTATTTC	CTCCATTATT	GTTAGATAAA	CCACAATTAA CTATAG	CTAC 7156
20	ACACTGAGCC A	GTAAGAGTA	CCAGGATG	СТТАСАВАТТ	GGCAATGCTT CAGAGG	AGAA 7216
	TTCCATCTCA T	GAAGACTCT	TTTTGAGTGG	AGATTTGCCA	ATAAATATCC GCTTTC	ATGC 7276
	CCACCCAGTC C	CCACTGAAA	GACAGTTAGG	ATATGACCTT	AGTGAAGGTA CCAAGG	GGCA 7336
	ACTTGGTAGG G	AGAAAAAAG	CCACTCTAAA	ATATAATCCA	AGTAAGAACA GTGCAT	ATGC 7396
	AACAGATACA G	CCCCCAGAC	AAATCCCTCA	GCTATCTCCC	TCCAACCAGA GTGCCA	CCCC 7456
	TTCACCTCAC A	ATTTGGAGT	CCCCATTCTA	GACCTGACAG	GCAGCTTAGT TATCAA	AATA 7516
25	CCATAAGAGG C	CTGGGATGG	AAGGGTAGGG	TGGAAAGGGT	TAAGCATGCT GTTACT	GAAC 7576
	AACATAATTA G	AAGGGAAGG	AGATGGCCAA	GCTCAAGCTA	TGTGGGATAG AGGAAA	ACTC 7636
	ACCTCCAGAG G	CAGATTCAG	AAACTGGGAT	AAGTCCGAAC	CTACAGGTGG ATTCTT	GTTG 7696
	AGGGAGACTG G	TGAAAATGT	TAAGAAGATG	GAAATAATGC	TTGGCACTTA GTAGGA	ACTG 7756
	GGCAAATCCA T	ATTTGGGGG	AGCCTGAAGT	TTATTCAATT	TTGATGGCCC TTTTAA	ATAA 7816
30	AAAGAATGTG G	CTGGGCGTG	GTGGCTCACA	CCTGTAATCC	CAGCACTTTG GGAGGC	CGAG 7876
	GGGGGCGGAT C	ACCTGAAGT	CAGGAGTTCA	AGACCAGCCT	GACCAACATG GAGAAA	CCCC 7936
	ATCTCTACTA A	AAATACAAA	ATTAGCTGGG	CGTGGTGGCA	TATGCCTGTA ATCCCA	GCTA 7996
	CTCGGGAGGC T	GAGGCAGGA	GAATCTTTTG	AACCCGGGAG	GCAGAGGTTG CGATGA	GCCT 8056
	AGATCGTGCC A	TTGCACTCC	AGCCTGGGCA	ACAAGAGCAA	AACTCGGTCT CAAAAA	AAAA 8116
	AAAAAAAAAG T	GAAATTAAC	CAAAGGCATT	AGCTTAATAA	TTTAATACTG TTTTTA	AGTA 8176
35	GGGCGGGGGG T	GGCTGGAAG	AGATCTGTGT	AAATGAGGGA	ATCTGACATT TAAGCT	TCAT 8236
	CAGCATCATA G	CAAATCTGC	TTCTGGAAGG	AACTCAATAA	ATATTAGTTG GAGGGG	GGGA 8296
	GAGAGTGAGG G	GTGGACTAG	GACCAGTTTT	AGCCCTTGTC	TTTAATCCCT TTTCCT	GCCA 8356
	CTAATAAGGA T	CTTAGCAGT	GGTTATAAAA	GTGGCCTAGG	TTCTAGATAA TAAGAT	'ACAA 8416
	CAGGCCAGGC A	CAGTGGCTC	ATGCCTATAA	TCCCAGCACT	TTGGGAGGGC AAGGCG	SAGTG 8476
40	TCTCACTTGA G	ATCAGGAGT	TCAAGACCAG	CCTGGCCAGC	ATGGCGATAC TCTGTC	TCTA 8536
	СТАААААААА Т	ACAAAAATT	AGCCAGGCAT	GGTGGCATGC	ACCTGTAATC CCAGCT	'ACTC 8596
	GTGAGCCTGA G	GCAGAAGAA	TCGCTTGAAA	CCAGGAGGTG	TAGGCTGCAG TGAGCT	GAGA 8656
	TCGCACCACT G	CACTCCAGC	CTGGGCGACA	GAATGAGACT	TTGTCTCAAA AAAAGA	AAAA 8716
	GATACAACAG G	CTACCCTTA	TGTGCTCACC	TTTCACTGTT	GATTACTAGC TATAAA	GTCC 8776
	TATAAAGTTC T	TTGGTCAAG	AACCTTGACA	ACACTAAGAG	GGATTTGCTT TGAGAG	GTTA 8836
45	CTGTCAGAGT C	TGTTTCATA	TATATACATA	TACATGTATA	TATGTATCTA TATCCA	GGCT 8896
	TGGCCAGGGT T	CCCTCAGAC	TTTCCAGTGC	ACTTGGGAGA	TGTTAGGTCA ATATCA	ACTT 8956
	TCCCTGGATT C	CAGATTCAAC	CCCTTCTGAT	GTAAAAAAAA	AAAAAAAAA GAAAGA	AATC 9016
	CCTTTCCCCT T	GGAGCACTC	AAGTTTCACC	AGGTGGGGCT	TTCCAAGTTG GGGGTT	CTCC 9076
	AAGGTCATTG G	GATTGCTTT	CACATCCATT	TGCTATGTAC	CTTCCCTATG ATGGCT	GGGA 9136
50	GTGGTCAACA T	CAAAACTAG	GAAAGCTACT	GCCCAAGGAT	GTCCTTACCT CTATTC	TGAA 9196
	ATGTGCAATA A	GTGTGATTA	AAGAGATTGC	CTGTTCTACC	TATCCACACT CTCGCT	TTCA 9256
	ACTGTAACTT T	CTTTTTTTC	TTTTTTTCTT	TTTTTCTTTT	TTTTTGAAAC GGAGTO	TCGC 9316

	TCTGTCGCCC AGGCTAGAGT GCAGTGGCAC GATCTCAGCT CACTGCAAGC TCTGCCTCCC	
	GGGTTCACGC CATTCTCCTG CCTCACCCTC CCAAGCAGCT GGGACTACAG GCGCCTGCCA	9376
_	CCATGCCCAG CTAATTTTTT CTATTTTTT CTATTTTTT CTATTTTTTTT	9436
5	ATGGTCTCGA TCTCCTGAAC TTCTCATCATCATCATCATCATCATCATCATCATCATCAT	9496
	ACAGGCGTGA GCCATCCCAC CCCCCTONS CCCCCCAAAG TGCTGGGATT	9556
	TGTAATGTTA CTAGACCTTT TCAACTTTC TGTAACTTTC TATACTGGTT CATCTTCCCC	9616
	ATTTCAGATT AGTTCCAAAT TCATGGGGGGGGGAT ATTTCTCATT TATACATTAG	9676
	GTAGACAGCT GCAGAAGTCG CTCCCALTAG	9736
10	GACCCACACT TGTTGATAA CAACAACTAGT TTATACTTTC ATCAACTTAG	9796
, 0	ATTGAGAAGT TGGAGATAAC CCCCTGAGGT CAAGAGTTAT GACTACTGAT TCCACAACTG	9856
	AAGGATGAAG AAATGCTATT TTAAMTTTGC CIGCCATCCA GAGTCTTTCA GGCATCTTTG	9916
	GAATCTGTGC TGCCATGACG CCANALTON AGGITTCTCT ATCAGTGCTT AGGATCATGG	9976
	GAATCTGTGC TGCCATGAGG CCAAAATTAA GTCCAAAACA TCTACTGGTT AGGATCATGG CATGGAAGAA CCTTAGGTGG TGCCCACATG TTCTCATGGA TCTACTGGTT CCAGGATTAA	10036
	CATGGAAGAA CCTTAGGTGG TGCCCACATG TTCTGATCCA TCTGCTGTT CCAGGATTAA GCACTAACAG GAAAAGTGCA GGCAGCACTA CCAGTTGCATA TCCTGCAAAA TAGACATGCT	10036
15	GCACTAACAG GAAAAGTGCA GGCAGCACTA CCAGTTGGAT AACCTGCAAAA TAGACATGCT CAAGTAATCT AACCATTTCT CACAAGGCCC TATTCTCTCA CTCACAAG ATTATAGTTT	
	CAAGTAATCT AACCATTTCT CACAAGGCCC TATTCTGTGA CTGAAACATA CAAGAATCTG	10156
	CATTTGGCCT TCTAAGGCAG GGCCCAGCCA AGGAGACCAT ATTCAGGACA GAAATCTG ACTACTATGG AACTGGAGTG CTTGGCAGGG AACACACAT ATTCAGGACA GAAATTCAAG	10216
	ACTACTATGG AACTGGAGTG CTTGGCAGGG AAGACAGAGT CAAGGACTGC CAACTGAGCC	10276
	AATACAGCAG GCTTACACAG GAACCCAGGG CCTAGCCCTA CAACAATTAT TGGGTCTATT	10336
	CACTGTAAGT TTTAATTTCA GGCTCCACTG AAAGAGTAAG CTAAGATTAT TGGGTCTATT TGTCTCTCTC ACAGTTGGCT CAGAAATGAG AACTGCTCACTG CTAAGATTCC TGGCACTTTC	10396
20	TGTCTCTCT ACAGTTGGCT CAGAAATGAG AACTGGTCAG GCCAGGCATG CTGCCACTTTC CCTGGAATCC CAGCACTTTG GGAGGCCGAA GTCCCACGGCATG GTGGCTTACA	10456
	CCTGGAATCC CAGCACTTG GGAGGCCGAA GTGGGAGGGT CACTTGAGGC CAGGAGTTCA GGACCAGCTT AGGCAACAAA GTGAGATACC CCCTGAGGG TACTTGAGGC CAGGAGTTCA	10516
	GGACCAGCTT AGGCAACAAA GTGAGATACC CCCTGACCCC TTCTCTACAA AAATAAATTT	10576
	TAAAAATTAG CCAAATGTGG TGGTGTATAC TTACAGTCCC AGCTACTCAG GAGGCTGAGG CAGGGGGATT GCTTGAGCCC AGGAATTCAA CCCTGCACTC	10636
	CAGGGGGATT GCTTGAGCCC AGGAATTCAA GGCTGCAGTG AGCTATGATT TCACCACTGC ACTTCTGGCT GGGCAACAGA GCGAGACCCT CTGTCAAACG	10696
25	ACTTCTGGCT GGGCAACAGA GCGAGACCCT GTCTCAAAGC AAAAAGAAAA AGAAACTAGA	10756
25	ACTAGCCTAA GTTTGTGGGA GGAGGTCATC ATCGTCTTTA GCCGTGAATA GTTATTATAG AGGACAGAAA TTGACATTAG CCCAAAAAGC TTGTCGGGTT	10816
	AGGACAGAAA TTGACATTAG CCCAAAAAGC TTGTGGTCTT TGCTGGAACT CTACTTAATC TTGAGCAAAT GTGGACACCA CTCAATGGGA CACCACACA	10876
	TTGAGCAAAT GTGGACACCA CTCAATGGGA GAGGAGAGAA GTAAGCTGTT TGATGTAATC GGGAAAACTA GAGGCCTGGA ACTGAATATG CATGGCAMAA GTAAGCTGTT TGATGTATAG	10936
	GGGAAAACTA GAGGCCTGGA ACTGAATATG CATCCCATGA CAGGGAGAAT AGGAGATTCG GAGTTAAGAA GGAGAGGAG TCAGTACTGC TCTTCACACAC	10996
	GAGTTAAGAA GGAGAGGAGG TCAGTACTGC TGTTCAGAGA TTTTTTTTAT GTAACTCTTG AGAAGCAAAA CTACTTTTGT TCTGTTTGGT AAMACAGTAG	11056
30	AGAAGCAAAA CTACTTTGT TCTGTTTGGT AATATACTTC AAAACAAACT TCATATATTC	11116
	ALGER DIGHT ("POWNER OF A SECOND	11176
	Clu MG AAT	11233
	Glu Met Asn	
	CCT CCT GAT AAC ATC AAG GAT ACA AAA AGT GAC ATC ATA TTC TTT CAG	
	Pro Pro Asp Asn Ile Lys Asp Thr Lys Ser Asp Ile Ile Phe Phe Glu	11281
<i>35</i>	95 110 File File Glu	
	AGA AGT GTC CCA GGA CAT GAT AAT AAG ATG CAA TTT GAA TCT TCA TCA Arg Ser Val Pro Gly His Asp Asp Lyo Mat Cl	
	Arg Ser Val Pro Gly His Asp Asn Lys Met Gln Phe Glu Ser Ser Ser	11329
	TIC CAN CON THE SILV SET SET SET	
	INC GAA GUA TAC TITT CIN COM MON OF THE	
	Tyr Glu Gly Tyr Phe Leu Ala Cys Glu Lys Glu Arg Asp Leu Phe Lys	11377
40	125 130 Leu Phe Lys	
	CIC ATT TTG AAA AAA CAC CAM CAA MAA AAA	
	Leu Ile Leu Lys Lys Glu Asp Glu Leu Gly Asp Arg Ser Ile Met Phe	1425
	140 145 Arg Ser Ile Met Phe	
	ACT GTT CAA AAC GAA GAC TAGCTATTAA AATTTCATCA	
45	Thr Val Gln Asn Glu Asp	1464
.5	155	

(18) INFORMATION FOR SEQ ID NO: 18:

(i)SEQUENCE CHARACTERISTICS:

- (A)LENGTH: 471 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: double

55

			(1) TOI	0000) Y : .	rinea	3.5									
		(<u>1</u> i	L)MOI	LECUI	LE TY	PE:	CDNA	A to	mRNA	A							
5		(vi	()	A)ORO		SM: n	E: nouse live										
		/ i s	/\FE	TURI													
0		(() ()	A) NAN 3) LOC	ME/KE	: NC	147	pept: 71 METI		S							
_		(x)	L)SE(QUENC	CE DE	ESCR	IPTIC	วท: ร	SEQ I	ID NO): 18	3:					
5	Asn													AAT Asn			48
										CAG				GAG Glu	GAT		96
20	-			20					25					30 CTG			144
														Leu			
?5	TAC Tyr	ATG Met 50	TAC Tyr	AAA Lys	GAC Asp	AGT Ser	GAA Glu 55	GTA Val	AGA Arg	GGA Gly	CTG Leu	GCT Ala 60	GTG Val	ACC Thr	CTC Leu	TCT Ser	192
		AAG												AAG Lys			240
30	TCC					GAT								ATA Ile			288
	GAT Asp	CTC Leu	ATA Ile	Phe	TTT	CAG Gln	AAA Lys	CGT Arg	Val	CCA	GGA Gly	CAC His	AAC Asn	AAG Lys	ATG	GAG Glu	336
35	TTT Phe	GAA Glu	TCT Ser	100 TCA Ser	CTG Leu	TAT Tyr	GAA Glu	Gly	105 CAC His	TTT Phe	CTT Leu	GCT Ala	Cys	110 CAA Gln	AAG Lys	GAA Glu	384
														AAT Asn			432
40								ACT Thr									471
) IN	FORM	ATIO	N FO		Q ID	NO:	19:								
15		(i	(A)LE B)TY	NGTH PE:	: 9 amin											
50		(i	i)MO	LECU	LE T	YPE:	pep	tide									
		(v)FRA	GMEN'	T TY	PE:	N-te	rmin	al f	ragm	ent						

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Asn Phe Gly Arg Leu His Cys Thr Thr 1

5

10

15

20

25

30

35

40

45

50

55

(20) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 8.5 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 110 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

(21) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

 Ser Val
 Lys
 Ser
 Glu
 Lys
 Ile
 Ser
 Thr
 Leu
 Ser
 Cys
 Glu
 Asn
 Lys
 Ile
 80

 Ile
 Ser
 Phe
 Lys
 Glu
 Met
 Asn
 Pro
 Pro
 Asn
 Asn
 Ile
 Lys
 Asp
 Thr
 Lys

 Ser
 Asp
 Ile
 Ile
 Phe
 Phe
 Gln
 Arg
 Ser
 Val
 Pro
 Gly
 His
 Asp
 Asn
 Lys
 Lys
 Lys
 Ile
 Asp
 Asn
 Lys
 Glu
 Asp
 Asp
 Lys
 Glu
 Ile
 <t

(22) INFORMATION FOR SEQ ID NO: 22:

5

10

15

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 20 25 Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 85 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 105 100 110 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150 155

(23) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 2.5 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 115 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

(24) INFORMATION FOR SEQ ID NO: 24:

5

10

15

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 20 25 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ser Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 120 125 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

- (25) INFORMATION FOR SEQ ID NO: 25:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 157 amino acids

- (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

5

30

35

40

45

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:
- Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 15 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ala Glu Asn Lys Ile 70 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 85 20 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 105 100 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 125 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 25 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150 145
 - (26) INFORMATION FOR SEQ ID NO: 26:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 - Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 20 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 45 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ala Glu Asn Lys Ile 75 70 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 110 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 125 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu

130 135 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 145 150 5 (27) INFORMATION FOR SEQ ID NO: 27: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 157 amino acids 10 (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27: Asn Phe Gly Arg Leu His Ala Thr Thr Ala Val Ile Arg Asn Ile Asn Asp Gln Val Leu Phe Val Asp Lys Arg Gln Pro Val Phe Glu Asp Met 20 25 Thr Asp Ile Asp Gln Ser Ala Ser Glu Pro Gln Thr Arg Leu Ile Ile Tyr Met Tyr Lys Asp Ser Glu Val Arg Gly Leu Ala Val Thr Leu Ser 55 25 Val Lys Asp Ser Lys Met Ser Thr Leu Ser Cys Lys Asn Lys Ile Ile 70 75 Ser Phe Glu Glu Met Asp Pro Pro Glu Asn Ile Asp Asp Ile Gln Ser 85 Asp Leu Ile Phe Phe Gln Lys Arg Val Pro Gly His Asn Lys Met Glu 100 30 105 110 Phe Glu Ser Ser Leu Tyr Glu Gly His Phe Leu Ala Cys Gln Lys Glu 120 125 Asp Asp Ala Phe Lys Leu Ile Leu Lys Lys Lys Asp Glu Asn Gly Asp 135 Lys Ser Val Met Phe Thr Leu Thr Asn Leu His Gln Ser 35 150 (28) INFORMATION FOR SEQ ID NO: 28: (i) SEQUENCE CHARACTERISTICS: 40 (A) LENGTH: 157 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28: Asn Phe Gly Arg Leu His Cys Thr Thr Ala Val Ile Arg Asn Ile Asn 10 50 Asp Gln Val Leu Phe Val Asp Lys Arg Gln Pro Val Phe Glu Asp Met 2.5 Thr Asp Ile Asp Gln Ser Ala Ser Glu Pro Gln Thr Arg Leu Ile Ile 40 Tyr Met Tyr Lys Asp Ser Glu Val Arg Gly Leu Ala Val Thr Leu Ser 55 55 Val Lys Asp Ser Lys Met Ser Thr Leu Ser Cys Lys Asn Lys Ile Ile

	65 Ser	Phe	Glu	Glu		70 Asp	Pro	Pro	Glu		75 Ile	Asp	Asp	Ile		80 Ser
5	Asp	Leu	Ile		85 Phe	Gln	Lys	Arg	Val	90 Pro	Gly	His	Asn	Lys 110	95 Met	Glu
	Phe	Glu	Ser 115	100 Ser	Leu	Tyr	Glu	Gly 120		Phe	Leu	Ala	Ser 125		rys	Glu
10	Asp	Asp 130		Phe	Lys	Leu	Ile 135		Lys	Lys	Lys	Asp 140		Asn	Gly	Asp
	Lys 145	Ser	Val	Met	Phe	Thr 150		Thr	Asn	Leu	His 155		Ser			
15																
20		•														
25																
30																
35																
<i>‡0</i>																
45																
50																

	SEQUENCE LISTING
	(1) GENERAL INFORMATION:
5	(i) APPLICANT: NAME: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO
	(ii) TITLE OF INVENTION: OSTEOCLASTGENIC INHIBITORY AGENT
10	(iii) NUMBER OF SEQUENCES: 28
	(iv) ADDRESS: (A) ADDRESSEE: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO
15	(B) STREET:2-3, 1-CHOME, SHIMOISHII (C) CITY: OKAYAMA (E) COUNTRY: JAPAN (F) POSTAL CODE (ZIP): 700
20	(v) COMPUTER READABLE FORM:(A) MEDIUM TYPE: Floppy disk(B) COMPUTER: IBM PC compatible(C) OPERATING SYSTEM: PC-DOS/MS-DOS
25	<pre>(vii) PRIOR APPLICATION DATA: (A1) APPLICATION NUMBER: JP 55,468/1997 (B1) FILING DATE: 25-FEB-1997</pre>
	(2) INFORMATION FOR SEQ ID NO: 1:
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (v) FRAGMENT TYPE: internal fragment
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
	Asn Asp Gln Val Leu Phe 1 5
	(3) INFORMATION FOR SEQ ID NO: 2:
10	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 6 amino acids(B) TYPE: amino acid(D) TOPOLOGY: linear
15	(ii) MOLECULE TYPE: internal fragment
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
	Phe Glu Asp Met Thr Asp 1 5
50	(4) INFORMATION FOR SEQ ID NO: 3:
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 7 amino acids
55	(B)TYPE: amino acid (D)TOPOLOGY: linear

		(ii)	MOLE	ECULE	TYF	E: b	epti	.de								
		(v) F	RAGM	MENT	TYPE	: ir	nterr	nal f	ragm	ment						
5		(xi)	SEQU	JENCE	DES	CRIE	OITS	1: SE	Q II	NO:	3:					
	Phe 1	Lys I	Leu 1	le I		ys I	Lys									
0	(5)	INFO	ORMAT	rion	FOR	SEQ	ID 8	10: 4	: :							
.5		(i)5	(A)	ENCE) LENC) TYP!) TOP(GTH: E: ar	5 ar mino	nino acio	acio d								
J		(ii) MOL	ECULI	E TY	PE: :	inte	rnal	frag	gment	5					
		(xi) SEQI	UENC	E DES	SCRI	PTIO	N: SI	EQ II	ои о	: 4:					
20	Met 1	Tyr	Lys .		Ser 5											
	(6)	INF	ORMA	TION	FOR	SEQ	ID	NO:	5 :							
25	(6) INFORMATION FOR SEQ ID NO: 5: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 5 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear															
		(ii) MOL	ECUL	E TY	PE:	inte	rnal	fra	gmen	t					
30		(xi) SEQ	UENC	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 5:					
	Ser 1	Thr	Leu	Ser	Cys 5											
35	(7)	INF	ORMA	TION	FOR	SEQ	ID	NO:	6 :							
40		(i)	(A (E	JENCE () LEN 3) TYF	IGTH: E: a	157 minc	ami aci	.no a .d		i						
	(D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide															
				QUENC					EO I	D NC): 6:					
45	Tvr	Phe											Arg	Asn	Leu	Asn
	1	Gln	_	-	5					10					15	
	_	Thr		20					25					30		
50		Ser	35					40					45			
		50 50 Val					55					60				
	6.5	Ser				70					75					80
5 <i>5</i>		: Sel : Asp			85					90					95	
	ser	. Asp	тте	TIG	FILE	E 11G	0111	A. g	J C 1	▼ CA ⊥	110	I				_1 ~

		a 1	 1	100					105					110			
			~					120					125			Glu	
5		T 2 0					135	Leu	Ile			140	Glu	Asp	Glu	Leu	
	Gly 145	Asp	Arg	Ser	Ile	Met 150	Phe	Thr	Val	Gln	Asn 155	Glu	Asp				
10	(8)	IN	FORM	ATIO	N FO	R SE	Q ID	NO:	7:								
10		(i	(.	UENC A) LE: B) TY D) TO	NGTH PE:	: 15 amin	7 am o ac	ino . id	S: acid	s							
15		(i	i) MO	LECU	LE T	YPE:	pep	tide									
		(×	i)SE	QUEN	CE D	ESCR	IPTI	: : MC	SEQ	ID N): 7	:					
	Asn 1	Phe	Gly	Arg	Leu 5	His	Cys	Thr	Thr		Val	Ile	Arg	Asn		Asn	
20		Gln	Val	Leu 20	Phe	Val	Asp	Lys	Arg	10 Gln	Pro	Val	Phe		15 Asp	Met	
	Thr	Asp	Ile 35		Gln	Ser	Ala	Ser	25 Glu	Pro	Gln	Thr	Arg	30 Leu	Ile	Ile	-
25	Tyr	Met 50	Tyr	Lys	Asp	Ser	Glu 55	40 Val	Arg	Gly	Leu		45 Val	Thr	Leu	Ser	
	Val 65		Asp	Ser	Lys	Met 70		Thr	Leu	Ser		60 Lys	Asn	Lys	Ile		
		Phe	Glu	Glu	Met 85		Pro	Pro	Glu		75 Ile	Asp	Asp	Ile	Gln	80 Ser	
30	Asp	Leu	Ile	Phe 100		Gln	Lys	Arg	Val	90 Pro	Gly	His	Asn		95 Met	Glu	
	Phe	Glu	Ser		Leu	Tyr	Glu	Gly	105 His	Phe	Leu	Ala		110 Gln	Lys	Glu	
	Asp	Asp	Ala	Phe	Lys	Leu	Ile	120 Leu	Lys	Lys	Lys		125 Glu	Asn	Gly	Asp	
35	Lys 145		Val	Met	Phe	Thr 150	135 Leu	Thr	Asn	Leu	His 155	140 Gln	Ser				
	(9)]	NFOR	TAMS	ON E	FOR S	EQ I	D NC): 8:									
40		(i)	(E) TYE	IGTH: PE: 11 LANDE	471 ucle DNES	. bas ic a S: d	e pa cid loubl	irs								
45		(ii	.) MOI	ECUL	E TY	PE:	CDNA										
		(vi		GINA) ORG) CEL	ANIS	M: h	uman										
50		(ix	(E) NAM) LOC	E/KE	N: 1	47	epti 1 METH		E							
66		(xi)SEQ	UENC	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 8:						
55	TAC	TTT	GGC	AAG	CTT	GAA	TCT	AAA	TTA	TCA	GTC	ATA	AGA	AAT	TTG	AAT	48

	Tyr	Pne	GIY	гÀг	Leu	GIU	Ser	Lys	Leu	ser 10	vai	lle	Arg	Asn	Leu 15	Asn	
	GAC	CAA	GTT	CTC	-	ATT	GAC	CAA	GGA		CGG	CCT	СТА	TTT		GAT	96
5			Val														,,
	ATG	ACT	GAT		GAC	TGT	AGA	GAT		GCA	CCC	CGG	ACC		TTT	ATT	144
			Asp 35														
	ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC	192
10			Met														
	TCT	GTG	AAG	TGT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	TGT	GAG	AAC	AAA	ATT	240
	65		Lys	_		70					75	_			-	80	
16			TTT														288
15			Phe		85					90			_	_	95	_	
			ATC														336
		_	Ile	100				_	105			-		110		-	22.
20			TTT					Tyr					Leu				384
	222	CAG	115 AGA	GAC	ميلي	بلىئىت	מממ	120 CTC	א יידיי	ጥጥር	מממ	222	125	CAT	CAA	TTC	432
			Arg														432
	-,0		30				135				-1-	140					
25			AGA														471
	Gly 145	Asp	Arg	Ser	Ile	Met 150	Phe	Thr	Val	Gln	Asn 155	Glu	Asp				
	(10) IN	FORM	ATIO	v FOI	R SEC	Q ID	NO:	9 :								
30		(i) SEQ	JENCI	E CHI	ARACT	reri:	STIC	S:								
			()	A) LEI B) TYI D) TOI	PΕ: 3	amino	o ac	id	cids								
		(i	i) MO:														
35									al f	r a crm	ant						
			FRA									_					
			i)SE	-								:					
40	Met 1	Tyr	Phe	GIY	Lys 5	Leu	Glu	Ser	Lys	Leu 10	Ser						
	(11) IN	FORM	ATIO	N FO	R SE	Q ID	NO:	10:								
45		(i) SEQ	UENC: A) LE													
			(B) TY D) TO	PE:	amin	o ac	id	-145								
		(i	i) MO	LECU	LE T	YPE:	pep	tide									
50		(v) FRA	GMEN	T TY	PE:	C-te	rmin	al f	ragm	ent						
		(x	i)SE	QUEN	CE D	ESCR	IPTI	ON:	SEQ	ID N	0: 1	0:					
55	Ser	Ile	Met	Phe	Thr 5	Val	Gln	Asn	Glu	Asp							
JJ	- .				_												

	(12) INFORMATION FOR SEQ ID NO: 11:
5	(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 13 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear
	(ii) MOLECULE TYPE: peptide
10	(v) FRAGMENT TYPE: N-terminal fragment
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
15	Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg 1 5 10
	(13) INFORMATION FOR SEQ ID NO: 12:
20	(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 14 amino acids (B)TYPE: amino acid (D)TOPOLOGY: linear
	(ii) MOLECULE TYPE: peptide
25	(v) FRAGMENT TYPE: internal fragment
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
	Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg
30	(14) INFORMATION FOR SEQ ID NO: 13:
<i>35</i>	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 17 amino acids(B) TYPE: amino acid(D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: peptide
	(v) FRAGMENT TYPE: internal fragment
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
	Ile Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 1 10 15
45	(15) INFORMATION FOR SEQ ID NO: 14:
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 471 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: CDNA
5 <i>5</i>	<pre>(ix) FEATURE: (A) NAME/KEY: mat peptide (B) LOCATION: 1471 (C) IDENTIFICATION METHOD: S</pre>

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

	TAC	TTT	GGC	AAG	CTT	GAA	TCT	AAA	TTA	TCA	GTC	ATA	AGA	AAT	TTG	AAT	48
5	Tyr 1	Phe	Gly	Lys	Leu 5	Glu	Ser	Lys	Leu	Ser 10	Val	Ile	Arg	Asn	Leu 15	Asn	
	GAC	CAA	GTT	CTC	TTC	ATT	GAC	CAA	GGA	AAT	CGG	CCT	CTA	TTT	GAA	GAT	96
	Asp	Gln	Val	Leu 20	Phe	Ile	Asp	Gln	Gly 25	Asn	Arg	Pro	Leu	Phe 30	Glu	Asp	
	ATG	ACT	GAT	TCT	GAC	TCT	AGA	GAT	AAT	GCA	CCC	CGG	ACC	ATA	TTT	ATT	144
10	Met	Thr	Asp 35	Ser	Asp	Ser	Arg	Asp 40	Asn	Ala	Pro	Arg	Thr 45	Ile	Phe	Ile	
	ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC	192
	Ile	Ser 50	Met	Tyr	Lys	Asp	Ser 55	Gln	Pro	Arg	Gly	Met 60	Ala	Val	Thr	Ile	
	TCT	GTG	AAG	TCT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	GCT	GAG	AAC	AAA	ATT	240
15	Ser 65	Val	Lys	Ser	Glu	Lys 70	Ile	Ser	Thr	Leu	Ser 75	Ala	Glu	Asn	Lys	Ile 80	
		TCC														AAA	288
	Ile	Ser	Phe	Lys	Glu 85	Met	Asn	Pro	Pro	Asp 90	Asn	Ile	Lys	Asp	Thr 95	Lys	
20	AGT	GAC	ATC	ATA	TTC	TTT	CAG	AGA	AGT	GTC	CCA	GGA	CAT	GAT	AAT	AAG	335
20	Ser	Asp	Ile	Ile 100	Phe	Phe	Gln	Arg	Ser 105	Val	Pro	Gly	His	Asp 110	Asn	Lys	
		CAA															384
		Gln	115					120		_	_		125		-		
25	AAA	GAG	AGA	GAC	CTT	TTT	AAA	CTC	ATT	TTG	AAA	AAA	GAG	GAT	GAA	TTG	432
		Glu 130					135				_	140		Asp	Glu	Leu	
		GAT															471
	-	Asp	Arg	Ser	Ile		Phe	Thr	Val	Gln		Glu	Asp				
30	145					150					155						
	(15) IN	FORM	OITA	1 FO	R SE	Q ID	NO:	15:								
		(i) SEQ														
					NGTH				cids								
35					PE: 8			ıd									

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: N-terminal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser 5

- (17) INFORMATION FOR SEQ ID NO: 16:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 471 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (ix) FEATURE:

40

45

50

- (A) NAME/KEY: mat peptide
- (B) LOCATION: 1..471

(C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

5	TAC Tyr 1	TTT Phe	GGC Gly	AAG Lys	CTT Leu 5	GAA Glu	TCT Ser	AAA Lys	TTA Leu	TCA Ser 10	GTC Val	ATA Ile	AGA Arg	AAT Asn	Leu	AAT Asn	48
	GAC	CAA Gln	GTT Val	Leu	TTC	ATT Ile	GAC Asp	CAA Gln	Gly	AAT	CGG Arg	CCT Pro	CTA Leu	Phe	15 GAA Glu	GAT Asp	96
10	ATG Met	ACT Thr	Asp	20 TCT Ser	GAC Asp	TCT Ser	AGA Arg	Asp	25 AAT Asn	GCA Ala	CCC Pro	CGG Arg	Thr	30 ATA Ile	TTT Phe	ATT Ile	144
	ATA Ile	AGT Ser	35 ATG Met	TAT Tyr	AAA Lys	GAT Asp	Ser	40 CAG Gln	CCT Pro	AGA Arg	GGT Gly	ATG Met	45 GCT Ala	GTA Val	ACT Thr	ATC Ile	192
15	Ser	50 GTG Val	AAG Lys	TCT Ser	GAG Glu	Lys	55 ATT Ile	TCA Ser	ACT Thr	CTC Leu	Ser	60 GCT Ala	GAG Glu	AAC Asn	AAA Lys	ATT Ile	240
20	65 ATT Ile	TCC Ser	TTT Phe	AAG Lys	Glu	70 ATG Met	AAT Asn	CCT Pro	CCT Pro	Asp	75 AAC Asn	ATC Ile	AAG Lys	GAT Asp	ACA Thr	80 AAA Lys	288
20	AGT Ser	GAC Asp	ATC Ile	ATA Ile	85 TTC Phe	TTT Phe	CAG Gln	AGA Arg	AGT Ser	90 GTC Val	CCA Pro	GGA Gly	CAT His	GAT Asp	95 AAT Asn	AAG Lys	336
25	ATG Met	CAA Gln	TTT Phe	100 GAA Glu	TCT Ser	TCA Ser	TCA Ser	TAC Tyr	105 GAA Glu	GGA Gly	TAC Tyr	TTT Phe	CTA Leu	110 GCT Ala	TCT Ser	GAA Glu	384
	AAA	GAG Glu	115 AGA	GAC	CTT	TTT	AAA	120 CTC	ATT	TTG	AAA	AAA	125 GAG	GAT	GAA	TTG	432
30	GGG	130 GAT Asp	AGA	TCT	ATA	ATG	135 TTC	ACT	GTT	CAA	AAC	140 GAA	GAC				471
	145	-	J			150					155						
	(18)	INFO	ORMAT	NOI	FOR	SEQ	ID N	10: 3	L7:								
35		(i)	(E	A) LEN		114 ucle	64 b	ase acid	pair	s							~
					POLOG	_											
40		(ii	L) MOI	LECUI	E TY	PE:	geno	omic	DNA								
		(vi		A) ORC	AL SC SANIS LL TY	M: h	umar		ì								
45		(i)	(E	A) NAM B) LOC C) IDE	ENTIF	N: 1	3	METH	HOD:								
50			E) () () ()	3) LOC 2) IDE 4) NAM 3) LOC	CATIC ENTIF ME/KE CATIC	N: 4 FICAT Y: i	82 TION .ntro .31	METF n L453	IOD:	S							
55			(<i>I</i>	A) NAN	ME/KE	Y: 1 N: 1	.eade .454 .	er pe	eptid	le							

5	(A)NAME/KEY: intron (B)LOCATION: 14654848 (C)IDENTIFICATION METHOD: E (A)NAME/KEY: leader peptide (B)LOCATION: 48494865	
	(C) IDENTIFICATION METHOD: S (A) NAME/KEY: mat peptide (B) LOCATION: 48664983 (C) IDENTIFICATION METHOD: S	
0	(A) NAME/KEY: intron (B) LOCATION: 49846317 (C) IDENTIFICATION METHOD: E (A) NAME/KEY: mat peptide	
5	(B)LOCATION: 63186451 (C)IDENTIFICATION METHOD: S (A)NAME/KEY: intron (B)LOCATION: 645211224	
20	(C)IDENTIFICATION METHOD: E (A)NAME/KEY: mat peptide (B)LOCATION: 1122511443 (C)IDENTIFICATION METHOD: S (A)NAME/KEY: 3'UTR	
	(B) LOCATION: 1144411464 (C) IDENTIFICATION METHOD: E (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:	
?5	AAG ATG GCT GCA CCA GTA GAA GAC AAT TGC ATC AAC TTT GTG GCA Met Ala Ala Glu Pro Val Glu Asp Asn Cys Ile Asn Phe Val Ala -35 -30 -25	48
30	ATG AAA TTT ATT GAC AAT ACG CTT TAC TTT ATA G GTAAGG CTAATGCCAT Met Lys Phe Ile Asp Asn Thr Leu Tyr Phe Ile Ala -20 -10	98
	AGAACAAATA CCAGGTTCAG ATAAATCTAT TCAATTAGAA AAGATGTTGT GAGGTGAACT ATTAAGTGAC TCTTTGTGTC ACCAAATTTC ACTGTAATAT TAATGGCTCT TAAAAAAATA GTGGACCTCT AGAAATTAAC CACAACATGT CCAAGGTCTC AGCACCTTGT CACACCACGT GTCCTGGCAC TTTAATCAGC AGTAGCTCAC TCTCCAGTTG GCAGTAAGTG CACATCATGA AAATCCCAGT TTTCATTGGA TTTCCATGGG AAAAATCCCA	158 218 278 338
35	GTACAAAACT GGGTGCATTC AGGAAATACA ATTTCCCAAA GCAAATTGGC AAATTATGTA AGAGATTCTC TAAATTTAGA GTTCCGTGAA TTACACCATT TTATGTAAAT ATGTTTGACA AGTAAAAATT GATTCTTTT TTTTTTTCT GTTGCCCAGG CTGGAGTGCA GTGGCACAAT CTCTGCTCAC TGCAACCTCC ACCTCCTGGG TTCAAGCAAT TCTCCTGCCT CAGCCTTCTG	398 458 518 578 638
±0	AGTAGCTGGG ACTACAGGTG CATCCCGCCA TGCCTGGCTA ATTTTTGGGT ATTTTTACTA GAGACAGGGT TTTGGCATGT TGTCCAGGCT GGTCTTGGAC TCCTGATCTC AGATGATCCT CCTGGCTCGG GCTCCCAAAG TGCTGGGATT ACAGGCATGA ACCACCACAC ATGGCCTAAA AATTGATTCT TATGATTAAT CTCCTGTGAA CAATTTGGCT TCATTTGAAA GTTTGCCTTC ATTTGAAACC TTCATTTAAA AGCCTGAGCA ACAAAGTGAG ACCCCATCTC TACAAAAAAC TGCAAAATAT CCTGTGGACA CCTCCTACCT TCTGTGGAGG CTGAAGCAGG AGGATCACTT	698 758 818 878 938 998
45	GAGCCTAGGA ATTTGAGCCT GCAGTGAGCT ATGATCCCAC CCCTACACTC CAGCCTGCAT GACAGTAGAC CCTGACACACA ACACACAAA AAAAACCTTC ATAAAAAATT ATTAGTTGAC TTTTCTTAGG TGACTTTCCG TTTAAGCAAT AAATTTAAAA GTAAAATCTC TAATTTTAGA AAATTTATTT TTAGTTACAT ATTGAAATTT TTAAACCCTA GGTTTAAGTT TTATGTCTAA	1058 1118 1178 1238
50	ATAATATTCT GATGAAAGCC AAGACAGACC CTTAAACCAT AAAAATAGGA GTTCGAGAAA GAGGAGTAGC AAAAGTAAAA GCTAGAATGA GATTGAATTC TGAGTCGAAA TACAAAATTT TACATATTCT GTTTCTCTCT TTTTCCCCCCT CTTAG CT GAA GAT GAT G GTAAA Ala Glu Asp Asp Glu	1418
55	GTAGAAATGA ATTTATTTTT CTTTGCAAAC TAAGTATCTG CTTGAGACAC ATCTATCTCA CCATTGTCAG CTGAGGAAAA AAAAAAATGG TTCTCATGCT ACCAATCTGC CTTCAAAGAA ATGTGGACTC AGTAGCACAG CTTTGGAATG AAGATGATCA TAAGAGATAC AAAGAAGAAC CTCTAGCAAA AGATGCTTCT CTATGCCTTA AAAAATTCTC CAGCTCTTAG AATCTACAAA	1590 1650

	ATAGACTTTG	CCTGTTTCAT	TGGTCCTAAG	ATTAGCATGA	AGCCATGGAT	TCTGTTGTAG	1770
	GGGGAGCGTT	GCATAGGAAA	AAGGGATTGA	AGCATTAGAA	TTGTCCAAAA	TCAGTAACAC	1830
	CTCCTCTCAG	AAATGCTTTG	GGAAGAAGCC	TGGAAGGTTC	CGGGTTGGTG	GTGGGGTGGG	1890
	GCAGAAAATT	CTGGAAGTAG	AGGAGATAGG	AATGGGTGGG	GCAAGAAGAC	CACATTCAGA	1950
5	GGCCAAAAGC	TGAAAGAAAC	CATGGCATTT	ATGATGAATT	CAGGGTAATT	CACAATCCAA	
	GTAGAGTAGG	AGTAGGAGAC	TGETGAGAGG	ACCTACACTC	ATAAACAGGG	TCTACACIGAA	2010
	CACCTTCTCT	CACCCCAACA	TOTICANATOT	CCACOMOTATIC	TTGGAGATAA	TGTAGAGCAA	2070
	TARCACATA	ATCTATTACC	TOCCCENTAC	DGACITIAIC	TIGGAGAIAA	TAGGGTTAAT	2130
	1AAGCACAA1	TCARTIAGC	DAAAIDDEAI	ATTAGTTTGT	TGTAACAAAG	ACATCCAAAG	2190
	ATACAGTAGC	TGAATAAGAT	AGAGAATTTT	TCTCTCAAAG	AAAGTCTAAG	TAGGCAGCTC	2250
10	AGAAGTAGTA	TGGCTGGAAG	CAACCTGATG	ATATTGGGAC	CCCCAACCTT	CTTCAGTCTT	2310
	GTACCCATCA	TCCCCTAGTT	GTTGATCTCA	CTCACATAGT	TGAAAATCAT	CATACTTCCT	2370
	GGGTTCATAT	CCCAGTTATC	AAGAAAGGGT	CAAGAGAAGT	CAGGCTCATT	CCTTTCAAAG	2430
	ACTCTAATTG	GAAGTTAAAC	ACATCAATCC	CCCTCATATT	CCATTGACTA	GAATTTAATC	2490
	ACATGGCCAC	ACCAAGTGCA	AGGAAATCTG	GAAAATATAA	TCTTTATTCC	AGGTAGCCAT	2550
	ATGACTCTTT	AAAATTCAGA	AATAATATAT	TTTTAAAATA	TCATTCTGGC	TTTGGTATAA	2610
15	AGAATTGATG	GTGTGGGGTG	AGGAGGCCAA	AATTAAGGGT	TGAGAGCCTA	TTATTTACT	2670
	TATTACAAGA	AATGATGGTG	TCATGAATTA	AGGTAGACAT	AGGGGAGTGC	TCATCACCAC	2730
	CTGTGAATGG	ATTTTAGAAA	CACTTGAGAG	DATCAATACC	ACATGATTTA	CCCTTCCATT	
	TGGAAAGGAG	AAGAAAGTAG	AAAAGATGAT	CCCTACATET	TTCACTTAGG	CARTICGALL	2790
	CATTCACTCA	AATACCCAAC	ACACCACCAA	CACCACCTT	TGGTGTATAC	CAATTIGTAC	2850
	ATCCATCACC	CATTAGGGAAC	TCCATCTCAC	ATCTCTCTCT	IGGIGIATAC	AAAGAGGAGG	2910
20	CACAAACTC	CATITCGITI	1GGA1C1GAG	AIGICIGIGG	AACGTCCTAG	TGGAGATGTC	2970
20	CACAAACICI	TCTACATGIG	GITCIGAGIT	CAGGACACAG	ATTTGGGCTG	GAGATAGAGA	3030
	TATTGTAGGC	Tratacatag	AAATGGCATT	TGAATCTATA	GAGATAAAAA	GACACATCAG	3090
	AGGAAATGTG	TAAAGTGAGA	GAGGAAAAGC	CAAGTACTGT	GCTGGGGGGA	ATACCTACAT	3150
	TTAAAGGATG	CAGTAGAAAG	AAGCTAATAA	ACAACAGAGA	GCAGACTAAC	CAAAAGGGGA	3210
	GAAGAAAAAC	CAAGAGAATT	CCACCGACTC	CCAGGAGAGC	ATTTCAAGAT	TGAGGGGATA	3270
25	GGTGTTGTGT	TGAATTTTGC	AGCCTTGAGA	ATCAAGGGCC	AGAACACAGC	TTTTAGATTT	3330
	AGCAACAAGG	AGTTTGGTGA	TCTCAGTGAA	AGCAGCTTGA	TGGTGAAATG	GAGGCAGAGG	3390
	CAGATTGCAA	TGAGTGAAAC	AGTGAATGGG	AAGTGAAGAA	ATGATACAGA	TAATTCTTGC	3450
	TAAAAGCTTG	GCTGTTAAAA	GGAGGAGAGA	AACAAGACTA	GCTGCAAAGT	GAGATTGGGT	3510
	TGATGGAGCA	GTTTTAAATC	TCAAAATAAA	GAGCTTTGTG	CTTTTTTGAT	TATGAAAATA	3570
	ATGTGTTAAT	TGTAACTAAT	TGAGGCAATG	AAAAAAGATA	ATAATATGAA	אמאמאטנוונ	3630
30	ATAAAAACCA	CCCAGAAATA	ATGATAGCTA	ССАТТТТСАТ	ACAATATTTC	TACATAAAAAT	3690
50	TCTATGTATA	TATACAGACA	CAGAAATGCT	ΤΑΤΑΤΤΤΤΤΑ	TTAAAAGGGA	TTCTTACTCCII	
	CCTAAGCTGC	$TTTTTTTT\Delta CT$	TAGTGATATA	TATGGACATC	TCTCCATGGC	A A CCA CTA A TI	3750
	TGCAGTTATA	TTAAGTTCAT	CATATTTCAC	AATAACCCCA	TATCTTTGCC	CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3810
	AATCAATTCT	TAATTCCTCA	VALVILLE CAC	CCACTTTCCT	GTTGTTATTA	CITITIATTT	3870
	CATAACCATT	CCTCTACACC	AIGITIGITI	CAGILIGII	GIIGIIAITA	ACAATGTTCC	3930
35	CCCACCACCA	ACARTACACC	COTTONIO	CATTIGICIG	ATTTTTTCTT	CAGGATAAAA	
33	CCCAGGAGGI	AGAATIGCIG	GGTTGATAGA	AGAGAAAGGA	TGATTGCCAA	ATTAAAGCTT	J4050
	CAGTAGAGGG	TACATGCCGA	GCACAAATGG	GATCAGCCCT	AGATACCAGA	AATGGCACTT	4110
	TCTCATTTCC	CCTTGGGACA	AAAGGGAGAG	AGGCAATAAC	TGTGCTGCCA	GAGTTAAATT	4170
	TGTACGTGGA	GTAGCAGGAA	ATCATTTGCT	GAAAATGAAA	ACAGAGATGA	TGTTGTAGAG	4230
	GTCCTGAAGA	GAGCAAAGAA	AATTTGAAAT	TGCGGCTATC	AGCTATGGAA	GAGAGTGCTG	4290
40	AACTGGAAAA	CAAAAGAAGT	ATTGACAATT	GGTATGCTTG	TAATGGCACC	GATTTGAACG	4350
40	CTTGTGCCAT	TGTTCACCAG	CAGCACTCAG	CAGCCAAGTT	TGGAGTTTTG	TAGCAGAAAG	4410
	ACAAATAAGT	TAGGGATTTA	ATATCCTGGC	CAAATGGTAG	ACAAAATGAA	CTCTGAGATC	4470
	CAGCTGCACA	GGGAAGGAAG	GGAAGACGGG	AAGAGGTTAG	ATAGGAAATA	CAAGAGTCAG	4530
	GAGACTGGAA	GATGTTGTGA	TATTTAAGAA	CACATAGAGT	TGGAGTAAAA	GTGTAAGAAA	4590
	ACTAGAAGGG	TAAGAGACCG	GTCAGAAAGT	AGGCTATTTG	AAGTTAACAC	TTCAGAGGCA	4650
45	GAGTAGTTCT	GAATGGTAAC	AAGAAATTGA	GTGTGCCTTT	GAGAGTAGGT	TAAAAAAACAA	4710
45	TAGGCAACTT	TATTGTAGCT	ACTTCTGGAA	CAGAAGATTG	TCATTAATAG	TOTOTOTOTO	4770
	СТАВАВТАТА	TAGCATACTT	ATTTGTCAAT	TAACAAAGAA	ACTATGTATT	TTTARGAMA	
	ארביים א מיניים ב	TATTCTAC	A AAC CTG C	אהטהמהטחת.	TAC TTT GGC	I I I AAA I GAG	4830
	ALLIAAIGII	CI	in han tou o	The Com Non	TAC III GGC	AAG CIT	4880
		د ی		ard ser Asp	Tyr Phe Gly	rla ren	
50	CAA mom are	, mm, max a	-5		T	5	
50	GAA TCT AAA	A TTA TCA GI	IC ATA AGA A	AAT TIG AAT	GAC CAA GTT	CTC TTC	4928
	Glu Ser Lys		at the Arg A		Asp Gln Val	Leu Phe	
		10		15		20	
	ATT GAC CAR	A GGA AAT CO	G CCT CTA T	TTT GAA GAT	ATG ACT GAT	TCT GAC	4976
	Ile Asp Glr	ı Gly Asn Ar	g Pro Leu B	Phe Glu Asp	Met Thr Asp	Ser Asp	
EE		25	3	3 0	35	-	
55	TGT AGA G	GTATTTTTT	TTAATTCGCA	AACATAGAAA	TGACTAGCTA	CTTCTTCCCA	5032

	Cys Arg Asp		
	40		
	TTCTGTTTTA CTGCTTACAT TGTTCCGTGC TAGTCCCAAT	CCTCAGATGA AAAGTCACAG	5092
	GAGTGACAAT AATTTCACTT ACAGGAAACT TTATAAGGCA	TCCACGTTTT TTAGTTGGGG	5152
5	TAAAAAATTG GATACAATAA GACATTGCTA GGGGTCATGC	CTCTCTGAGC CTGCCTTTGA	5212
	ATCACCAATC CCTTTATTGT GATTGCATTA ACTGTTTAAA	ACCTCTATAG TTGGATGCTT	5272
	AATCCCTGCT TGTTACAGCT GAAAATGCTG ATAGTTTACC	AGGTGTGGTG GCATCTATCT	5332
	GTAATCCTAG CTACTTGGGA GGCTCAAGCA GGAGGATTGC		5392
	CTGTAGTACA CTGTGATCGT ACCTGTGAAT AGCCACTGCA		5452
	AGACCTTGTC TCTAAAATTA AAAAAAAAA AAAAAAAAC	_	5512
10	AAGTCTACTG TGCCTTCCAA AACATGAATT CCAAATATCA		5572
	AGTGAATGTG CATTCTTTAA AAATACTGAA TACTTACCTT		5632
	TATTTAGCAT TTAAAAGTTA AAAACAATCT TTTAGAATTC		5692
	AAAGTTGCAG CGTGTGTGTT GTAATACACA TTAAACTGTG		5752
	ATGCAGTTTC ACTCTGTCAC CCAGGCTGAA GTGCAGTGCA		5812
15	CTCACTACAA CCTCCACCTC CCACGTTCAA GCGATTCTCA		5872
	GTGGGATTAC AGGCATGCAC CACTTACACC CGGCTAATTT		5932
	GGGTTTCACC ATGTTGGCCA GGCTGGTCTC AAACCCCTAA		5992
	TCAGCCTCCC AAACAAACAA ACAACCCCAC AGTTTAATAT		6052
	CAACTTTTAT GAGTATTTTA ATGATATAGA TTATAAAAGG		6112
	CTGGGATTAC AGGCATGAGC CACTGTGCCA GGCCTGAACT		6172
20	CCAGCTGTAC ATAGTCTCCT GCAGACTGGC CAAGTCTCAA		6232
	GGACTATCCT TTGGTTAAAT TTCCGCAAAT GTTCCTGTGC		6292
	TTCTCATTTA TTATATTTAT TTCAG AT AAT GCA CCC		6343
		Arg Thr Ile Phe Ile	
	40	45	
0.5	ATA AGT ATG TAT AAA GAT AGC CAG CCT AGA GGT	ATG GCT GTA ACT ATC	6391
25	Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly		
	50 55	60	
	TCT GTG AAG TGT GAG AAA ATT TCA ACT CTC TCC	TGT GAG AAC AAA ATT	6439
	Ser Val Lys Cys Glu Lys Ile Ser Thr Leu Ser		
	65 70 75	80	
30	ATT TCC TTT AAG GTAAG ACTGAGCCTT ACTTTGTTTT	CAATCATGTT AATATAATCA	6496
			0 2 0 0
	Ile Ser Phe Lys		0129
	Ile Ser Phe Lys ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA		6556
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT	
,	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT	6556
\$	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT	6556 6616
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA	6556 6616 6676
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT	6556 6616 6676 6736 6796 6856
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC	6556 6616 6676 6736 6796
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT	6556 6616 6676 6736 6796 6856 6916
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTAATGT TAATGCTATA ATTATCTTCA	6556 6616 6676 6736 6796 6856 6916 6976 7036
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATTCT	6556 6616 6676 6736 6796 6856 6916 6976 7036 7096
35	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC	6556 6616 6676 6736 6796 6856 6916 6976 7036 7096 7156
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TCCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216 7276
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCAAGGGGCA	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216 7276 7336
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCCACTCTAAA ATATAATCCA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGACA GTGCATATGC	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216 7336 7396
	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCCACTCTAAA ATATAATCCA AACAGATACA GCCCCCAGAC AAATCCCTCA GCTATCTCCC	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCAAGGGGCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216 7336 7396 7456
40	ATATAATTAG AAATATAACA TTATTTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC TGGCCTCTAC TTTTTCTCTA CTATTATTTT TCTCTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCCACTCTAAA ATATAATCCA AACAGATACA GCCCCCAGAC AAATCCCTCA GCTATCTCCC TTCAGGTGAC AATTTGGAGT CCCCCATTCTA GACCTGACAG	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCCAGGGGCA AGTAAGAACA GTGCATCCC GCAGCTTAGT TATCAAAATA	6556 6616 6676 6736 6796 6856 6916 7036 7096 7156 7216 7336 7456 7516
40	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA ACTTGGTAG GAGAAAAAAG CCACTCTAAA ATATAATCCA AACAGATACA GCCCCCAGAC AAATCCCTCA GCTATCTCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATGG AAGGGTAGGG TGGAAAGGGT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCAAGGGCCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC	6556 6616 6676 6736 6796 6856 6916 7036 7156 7216 7336 7336 7456 7516
40	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCA AACAGATACA GCCCCCAGC AAATCCCTCA GCTATCTCCC TTCAGGTGAC AATTTGGAGT CCCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATGG AAGGGTAGGG TGGAAAGGGT AACATAATTA GAAGGGAAGG AGATGGCCAA GCTCAAGCTA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTTCA TTATTCTCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCAAGGGCCA AGTAAGAACA GTGCATCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TGTGGGATAG AGGAAAACTC	6556 6616 6676 6736 6796 6856 6916 7096 7156 7216 7336 7456 7516 7576 7636
40	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTC TTTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCA AACAGATACA GCCCCCAGC AAATCCCTCA TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATG AAGGGTAGGG TGGAAAGGGT AACATAATTA GAAGGGAAGG AGATGGCCAA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGTA CCAAGGGCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TGTGGGATAG AGGAAAACTC CTACAGGTGG ATTCTTTTG	6556 6616 6676 6736 6796 6856 6976 7096 7216 7276 73396 7456 7576 7636 7696
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGTGATG CCACCCAGTC CCCACTGAAA GACAGTTAGG AGATTTGCCA ACTTGGTAG GAGAAAAAAG CACAGTTAGG ATATGACCTT ACTTGGTAG GAGAAAAAAG CCACTCTAAA ATATAATCCA AACAGATACA GCCCCAGAC AAATCCCTCA TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATG AAGTGGCCAA GCTCAAGCGT ACCTGAGAG GCAGATTCAG AGGGTAGGG TGGAAAGGGT AACATAATTA GAAGGGAAGG AGATGGCCAA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATG GAAATAATGC	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA AGTAAGAACA GTGCATATGC AGTGAAGGTA CCAAGGGCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAATA TAAGCATGCT GTTACTGAAC TGTGGGATAG AGGAAAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG	6556 6676 66736 6736 6736 6856 6976 7096 7216 7273 73396 74516 7576 7636 7756
40	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGTGATG TTCCATGTCA GTATATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGATG CTTACAAATT TCCATGTAG GAGAAAAAAG CCACTCTAAA ATATAATCCA ACAGATACA GCCCCAGAC AAATCCCTCA GCTATCTCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATG AAATCCCTCA GCTATCTCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATG AAGTGGCCAA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATG GAAATAATGC GCCAAATCCA TATTTGGGGG AGCCTGAAGT TTATTCAATT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTATA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGACA GTGCAACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTAAACC TCTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG TTGGCACTTA GTAGGAACTG TTGATGGCCC TTTTAAATAA	6556 6676 6736 6736 6736 6856 6976 7096 7216 7273 73396 7576 7536 7696 7756 7816
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGTGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCT ACTTGGTAGG GAGAAAAAAG CCCCCAGAC AAATCCCTCA ACAGATACA GCCCCAGAC AAATCCCTCA GCTATCCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GCATAAGAGG CCTGGGATG AAGTGGCCAA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATG GAAATAATGC GGCAAATCCA TATTTGGGGG AGCCTGAAGT TTATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTTGTAATCC	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGATA GCAAGGGGCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TGTGGGATAG AGGAAAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG TTGATGGCCC TTTTAAATAA CAGCACTTTG GGAGGCCGAG	6556 6676 66736 6736 6736 6856 6976 7096 7276 73396 7276 73396 7576 7696 77536 7816 7876
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGTGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA ACAGATACA GCCCCAGAC AAATCCCTCA ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTACCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATACCCC ACTTGGTAG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATCCC CCACCAGAC AAATCCCTCA GCTATCTCCC GCATAAGAGG CCTGGGATGG AAGCGTAGGG TGGAAAGGGT ACATAATTA GAAGGGAAGG AAATCCCTCA GCTATCCCC ACTGCAGAG CCTGGGATG AAATCCCTCA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATG GAAATAATGC GGCAAATCCA TATTTGGGGG AGCCTGAAGT TTATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC GGGGGCGGAT CACCTGAAGT CAGGGGTTCA AGACCAGCCT	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTATA ATTATCTTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTCCATC AGTGAAGGATA CCAAGGGGCA AGTAAGAACA GTGCATATGC TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TCTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG TTGATGGCCC TTTTAAATAA CAGCACTTTG GGAGGCCGAG GACCAACATG GAGAAACCCC	6556 6676 66736 6736 6736 6736 6976 7096 7276 73396 7276 73396 7576 7636 77536 7816 7836
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTTGAGTGG AGATTTGCCA ACATGGTAGA GCCCCAGAC ACACGGTTAGA ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG AACAGATACA GCCCCAGAC AAATCCCTCA GCTATCTCC CCACCAGG CATTGGAGT AAGTGGCCAA GCTCAAGGGT ACATAATTA GAAGGGAAGG AAGTGGCCAA GCTCAAGCTA AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGGGAGACTG GTGAAAATGT TAAGAAGATG GAAATAATGC GGCAAATCCA TATTTGGGGG AGCCTGAAGT TTATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC GGGGGGGGAT CACCTGAAGT CAGGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGAC GCAGCTTAGT TCCAACCAGA GTGCCACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TCTACAGGTGG ATTCTTGATC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTC CTACAGCTGG ATTCTTGTTG TTGGCACTTT GGGAGGCCGAG GACCAACATG GAGAAACCCC TATGCCTGTA ATCCCAGCTA	6556 6676 6736 6736 6736 6736 6976 7096 7276 7336 74516 75576 7636 77536 77536 77536 77536 77536 77536 77536 77536
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA ACACGATACA GCCCCCAGAC ACACCTTAAA ATATAATCCC CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATCCC CTAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG AACAGATACA GCCCCCAGC AAATCCCTCA GCTATCCCC GCATAAGAGG CCTGGGATGG AAGGGTAGGG TGGAAAGGGT AACATAATTA GAAGGGAAGG AAACTCGGAT AAGTCCGAAC AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATT GAAAGAATATCCGAAC AGGGAGACTG GTGAAAATGT TAAGAAGATT GAAAATACCAA AGACTGCGAGT CACCTGAAGT TTATTCAATT AAAAAAATCCAAA ATTATCCAAC AGCGGGGGGGAT CACCTGAAGT TAATTCCGACC AGGGGGGGGAT CACCTGAAGT TAAGAAGATT TAAGAAGATT TAAGAAGATT TAATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC GGGGGGGGGAT CACCTGAAGT CAGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA CTCGGGAGGC TGAGGATTCA AAAATACAAA ATTAGCTGGG CGTGGTGGCA ACCTGGAGG GAAATACAAA ATTAGCTGGG CGTGGTGGCA AAAATACAAA ATTAGCTGGG CGTGGTGGCA CTCGGGAGGC TGAGGATTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGACA GTGCAACCCC GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TCTACAGGTGG ATTCTTGATC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTC CTACAGCTGG ATTCTTGTTG TTGATGGCCC TTTTAAATAA CAGCACTTTG GGAGGCCGAG GACCAACATG GAGAAACCCC TATGCCTGTA ATCCCAGCTA GCAGAGGTTG CGATGAGCCT	6556 6676 6736 6736 6736 6736 6976 7096 7276 73396 7276 73396 75576 7636 77536 77536 7876 7876 7876 7876 7876 7876 7876 78
40 45	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATCCC CTATGATACA GCCCCCAGAC AAATCCCTCA GCTATCCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG GAAAAAAAAG CCACTCTAAA ATATAATCCC GCATAAGAGG CCTGGGATG AAGGGTAGG TGGAAAGGGT AACATAATTA GAAGGAAGG AAACTCGGAT AAGTCCGAAC AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATG GAAAAAATGC GGCAAATCCA TATTTGGGGG AGCCTGAAGT TTATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC GGGGGGGGGAT CACCTGAAGT CAGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA ACCTGGAGGC TGAGGATTCA AAACTACAAA ATTATCCAACTC CCCGGGAGGC TGAAAATGT TAAGAAGATG GAAAATAATGC GGGGGGGGGAT CACCTGAAGT CAGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA ACCTGGAGGC GTGAAAATGT TAATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC CCGGGGGGGGT CACCTGAAGT CAGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA AGATCGTGCC ATTGCACTCC AGCCTGGGCA ACAAGAGCAA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGAC GCAGCTTAGT TAGCAGCAGG GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TCTACAGGTGG ATTCTTGTTG TTGGGGATAG AGGAAAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG TTGATGGCCC TTTTAAATAA CAGCACTTG GGAGGCCGAG GACCAACATG GAGAAACCCC TATGCCTGTA ATCCCAGCTA GCAGAGGTTG CGATGAGCCT AACTCGGTTC CAAAAAAAAA	6556 6676 6736 6736 6736 6736 6976 7096 7276 73396 74516 75576 7696 77516 7876 7876 7876 7876 7876 7876 7876 78
40 45 50	ATATAATTAG AAATATAACA TTATTCTAA TGTTAATATA CAAATATCCT CAGACCAACC TTTTGTCTAG AACAGAAATA AAAGTGAATA CTTACTAAAA ATTATCAAAC TCTTTACCTA CTGAGCCTGT CACAGGGGAA GAGGAGATAC AACACTTGTT ACAATCAGTC TTTATACAAA TAATAATGTA GAATACATAT TAACATGTGA CTTTCCAGAA TGAGTTCTGC TATGAAGAAT ATTTCTACAC CTTTGTAAAT TATGATAATA TTTTAATCCC CTTAGCCTAA GTCTTAGACA CAAGCTTCAG CTTCCAGTTG TAATCTAATT GAATAAAAGT TATGAGATCA GCTGTAAAAG AGCCAGGTAT AAAGTATTC CTCCATTATT GTTAGATAAA AGACTGAGCC AGTAAGAGTA GCCAGGGATG CTTACAAATT TTCCATGTCA TGAAGACTCT TTTTGAGTGG AGATTTGCCA ACACGATACA GCCCCCAGAC ACACCTTAAA ATATAATCCC CCACCCAGTC CCCACTGAAA GACAGTTAGG ATATGACCTT ACTTGGTAGG GAGAAAAAAG CCACTCTAAA ATATAATCCC TTCAGGTGAC AATTTGGAGT CCCCATTCTA GCTATCCC CTAGGTGAC AATTTGGAGT CCCCATTCTA GACCTGACAG AACAGATACA GCCCCCAGC AAATCCCTCA GCTATCCCC GCATAAGAGG CCTGGGATGG AAGGGTAGGG TGGAAAGGGT AACATAATTA GAAGGGAAGG AAACTCGGAT AAGTCCGAAC AGCTGCAGAG GCAGATTCAG AAACTGGGAT AAGTCCGAAC AGCGGAGACTG GTGAAAATGT TAAGAAGATT GAAAGAATATCCGAAC AGGGAGACTG GTGAAAATGT TAAGAAGATT GAAAATACCAA AGACTGCGAGT CACCTGAAGT TTATTCAATT AAAAAAATCCAAA ATTATCCAAC AGCGGGGGGGAT CACCTGAAGT TAATTCCGACC AGGGGGGGGAT CACCTGAAGT TAAGAAGATT TAAGAAGATT TAAGAAGATT TAATTCAATT AAAGAATGTG GCTGGGCGTG GTGGCTCACA CCTGTAATCC GGGGGGGGGAT CACCTGAAGT CAGGAGTTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA CTCGGGAGGC TGAGGATTCA AAAATACAAA ATTAGCTGGG CGTGGTGGCA ACCTGGAGG GAAATACAAA ATTAGCTGGG CGTGGTGGCA AAAATACAAA ATTAGCTGGG CGTGGTGGCA CTCGGGAGGC TGAGGATTCA AGACCAGCCT ATCTCTACTA AAAATACAAA ATTAGCTGGG CGTGGTGGCA	AGTAATGTAA TTAGAAAACT ACAAGAAGCA GAGAACCATT TTGTGATAAT GATGGTTTTT TTATGACCTG CATCTCCTGA GTGAGTTATA CATTTAAGAA GAAGCTAATT ATCCTTCTAT TAGTTGTTTT GTTGCTGATC ATGTATGTTA TTTTTAATGT TAATGCTATA ATTATCTCA TTATTCTCCA TTATTATCT CCACAATTAA CTATAGCTAC GGCAATGCTT CAGAGGAGAA ATAAATATCC GCTTTCATGC AGTGAAGGAC GCAGCTTAGT TAGCAGCAGG GCAGCTTAGT TATCAAAATA TAAGCATGCT GTTACTGAAC TCTACAGGTGG ATTCTTGTTG TTGGGGATAG AGGAAAACTC CTACAGGTGG ATTCTTGTTG TTGGCACTTA GTAGGAACTG TTGATGGCCC TTTTAAATAA CAGCACTTG GGAGGCCGAG GACCAACATG GAGAAACCCC TATGCCTGTA ATCCCAGCTA GCAGAGGTTG CGATGAGCCT AACTCGGTTC CAAAAAAAAA	6556 6676 6736 6736 6736 6736 6976 7096 7276 73396 7276 73396 75576 7636 77536 77536 7876 7876 7876 7876 7876 7876 7876 78

	GGGCGGGGG	TGGCTGGAAG	AGATCTGTGT	AAATGAGGGA	ATCTGACATT	TAAGCTTCAT	8236
	CAGCATCATA	GCAAATCTGC	TTCTGGAAGG	AACTCAATAA	ATATTAGTTG	GAGGGGGGGA	8296
	GAGAGTGAGG	GGTGGACTAG	GACCAGTTTT	AGCCCTTGTC	TTTAATCCCT	TTTCCTGCCA	8356
	CTAATAAGGA	TCTTAGCAGT	GGTTATAAAA	GTGGCCTAGG	TTCTAGATAA	TAAGATACAA	8416
5	CAGGCCAGGC	ACAGTGGCTC	ATGCCTATAA	TCCCAGCACT	TTGGGAGGGC	AAGGCGAGTG	8476
	TCTCACTTGA	GATCAGGAGT	TCAAGACCAG	CCTGGCCAGC	ATGGCGATAC	TCTGTCTCTA	8536
	CTAAAAAAAA	TACAAAAATT	AGCCAGGCAT	GGTGGCATGC	ACCTGTAATC	CCAGCTACTC	8596
	GTGAGCCTGA	GGCAGAAGAA	TCGCTTGAAA	CCAGGAGGTG	TAGGCTGCAG	TGAGCTGAGA	8656
	TCGCACCACT	GCACTCCAGC	CTGGGCGACA	GAATGAGACT	TTGTCTCAAA	444644444	8716
10	GATACAACAG	GCTACCCTTA	TGTGCTCACC	TTTCACTGTT	GATTACTAGC	TATAAAGTCC	8776
. 3	TATAAAGTTC	TTTGGTCAAG	AACCTTGACA	ACACTAAGAG	GGATTTGCTT	TGAGAGGTTA	8836
	CTGTCAGAGT	CTGTTTCATA	TATATACATA	TACATGTATA	TATGTATCTA	TATCCAGGCT	8896
	TGGCCAGGGT	TCCCTCAGAC	TTTCCAGTGC	ACTTGGGAGA	TGTTAGGTCA	ATATCAACTT	8956
	TCCCTGGATT	CAGATTCAAC	CCCTTCTGAT	GTAAAAAAA	AAAAAAAAA	GAAAGAAATC	9016
	CCTTTCCCCT	TGGAGCACTC	AAGTTTCACC	AGGTGGGGCT	TTCCAAGTTG	GGGGTTCTCC	9076
15	AAGGTCATTG	GGATTGCTTT	CACATCCATT	TGCTATGTAC	CTTCCCTATG	ATGGCTGGGA	9136
	GTGGTCAACA	TCAAAACTAG	GAAAGCTACT	GCCCAAGGAT	GTCCTTACCT	CTATTCTGAA	9196
	ATGTGCAATA	AGTGTGATTA	AAGAGATTGC	CTGTTCTACC	TATCCACACT	CTCGCTTTCA	9256
	ACTGTAACTT	TCTTTTTTTC	TTTTTTTTTTT	TTTTTCTTTT	TTTTTGAAAC	GGAGTCTCGC	9316
	TCTGTCGCCC	AGGCTAGAGT	GCAGTGGCAC	GATCTCAGCT	CACTGCAAGC	TCTGCCTCCC	9376
22	GGGTTCACGC	CATTCTCCTG	CCTCACCCTC	CCAAGCAGCT	GGGACTACAG	GCGCCTGCCA	9436
20	CCATGCCCAG	CTAATTTTTT	GTATTTTTAG	TAGAGACGGG	GTTTCACCGT	GTTAGCCAGG	9496
	ATGGTCTCGA	TCTCCTGAAC	TTGTGATCCG	CCCGCCTCAG	CCTCCCAAAG	TGCTGGGATT	9556
	ACAGGCGTGA	GCCATCGCAC	CCGGCTCAAC	TGTAACTTTC	TATACTGGTT	CATCTTCCCC	9616
	TGTAATGTTA	CTAGAGCTTT	TGAAGTTTTG	GCTATGGATT	ATTTCTCATT	TATACATTAG	9676
	ATTTCAGATT	AGTTCCAAAT	TGATGCCCAC	AGCTTAGGGT	CTCTTCCTAA	ATTGTATATT	9736
25	GTAGACAGCT	GCAGAAGTGG	GTGCCAATAG	GGGAACTAGT	TTATACTTTC	ATCAACTTAG	9796
	GACCCACACT	TGTTGATAAA	GAACAAAGGT	CAAGAGTTAT	GACTACTGAT	TCCACAACTG	9856
	ATTGAGAAGT	TGGAGATAAC	CCCGTGACCT	CTGCCATCCA	GAGTCTTTCA	GGCATCTTTG	9916
	AAGGATGAAG	AAATGCTATT	TTAATTTTGG	AGGTTTCTCT	ATCAGTGCTT	AGGATCATGG	9976
	GAATCTGTGC	TGCCATGAGG	CCAAAATTAA	GTCCAAAACA	TCTACTGGTT	CCAGGATTAA	10036
	CATGGAAGAA	CCTTAGGTGG	TGCCCACATG	TTCTGATCCA	TCCTGCAAAA	TAGACATGCT	10096
30	GCACTAACAG	GAAAAGTGCA	GGCAGCACTA	CCAGTTGGAT	AACCTGCAAG	ATTATAGTTT	10156
	CAAGTAATCT	AACCATTTCT	CACAAGGCCC	TATTCTGTGA	CTGAAACATA	CAAGAATCTG	10216
	CATTTGGCCT	TCTAAGGCAG	GGCCCAGCCA	AGGAGACCAT	ATTCAGGACA	GAAATTCAAG	10276
	ACTACTATGG	AACTGGAGTG	CTTGGCAGGG	AAGACAGAGT	CAAGGACTGC	CAACTGAGCC	10336
	AATACAGCAG	GCTTACACAG	GAACCCAGGG	CCTAGCCCTA	CAACAATTAT	TGGGTCTATT	10396
35	CACTGTAAGT	TTTAATTTCA	GGCTCCACTG	AAAGAGTAAG	CTAAGATTCC	TGGCACTTTC	10456
55	TGTCTCTCTC	ACAGTTGGCT	CAGAAATGAG	AACTGGTCAG	GCCAGGCATG	GTGGCTTACA	10516
	CCTGGAATCC	CAGCACTTTG	GGAGGCCGAA	GTGGGAGGGT	CACTTGAGGC	CAGGAGTTCA	10576
	GGACCAGCTT	AGGCAACAAA	GTGAGATACC	CCCTGACCCC	TTCTCTACAA	TTTAAATAAA	10636
	TAAAAATTAG	CCAAATGTGG	TGGTGTATAC	TTACAGTCCC	AGCTACTCAG	GAGGCTGAGG	10696
	CAGGGGGATT	GCTTGAGCCC	AGGAATTCAA	GGCTGCAGTG	AGCTATGATT	TCACCACTGC	10756
40	ACTTCTGGCT	GGGCAACAGA	GCGAGACCCT	GTCTCAAAGC	AAAAAGAAAA	AGAAACTAGA	10816
	ACTAGCCTAA	GTTTGTGGGA	GGAGGTCATC	ATCGTCTTTA	GCCGTGAATG	GTTATTATAG	10876
	AGGACAGAAA	TTGACATTAG	CCCAAAAAGC	TTGTGGTCTT	TGCTGGAACT	CTACTTAATC	10936
	TTGAGCAAAT	GTGGACACCA	CTCAATGGGA	GAGGAGAA	GTAAGCTGTT	TGATGTATAG	10996
	GGGAAAACTA	GAGGCCTGGA	ACTGAATATG	CATCCCATGA	CAGGGAGAAT	AGGAGATTCG	11056
45	GAGTTAAGAA	GGAGAGGAGG	TCAGTACTGC	TGTTCAGAGA	TTTTTTTAT	GTAACTCTTG	11116
45	AGAAGCAAAA	CTACTTTGT	TCTGTTTGGT	AATATACTTC	AAAACAAACT	TCATATATTC	
	AAATTGTTCA	TGTCCTGAAA	TAATTAGGTA	ATGTTTTTT	CTCTATAG GA		11233
				•		u Met Asn	
	CCT CCT CA		C CAM ACA 1		85		
					ATC ATA TTC		11281
50		ASR ITE LY		lys ser Asp	Ile Ile Phe	Phe Glu	
	90	ם ממא ממא מי	95 • • • • • • • • • • • • • • • • • • •	אר איים מאא	100	mas mas	
	AGA AGI GIG	L CCA GGA CA	I GAI AAI A	AAG AIG CAA	TTT GAA TCT	TCA TCA	11329
	105	r bro Già Hi	s asp asn i 110	has wer GIU	Phe Glu Ser	ser ser	
		י אור הההה כם		מא אא מאם	115 AGA GAC CTT	מאל ווייטייט	11122
	TVY Glu Gl	r Tur Dhe Te	A GCI IGI (In the Clu	Arg Asp Leu	Dhe Tue	11377
55	120	TYP PRE LE		130	ura wab ren	Pne Lys 135	
	-£-U	1.4		130		132	

	CTC ATT TTG AAA AAA GAG GAT GAA TTG GGG GAT AGA TCT ATA ATG TTC Leu Ile Leu Lys Lys Glu Asp Glu Leu Gly Asp Arg Ser Ile Met Phe 140 145 150	425													
5	ACT CTT CAA AAC CAA CAC TACCTATTAA AATTTCATCATC	464													
	(19) INFORMATION FOR SEQ ID NO: 18:														
10	(i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 471 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: double (D)TOPOLOGY: linear														
15	(ii) MOLECULE TYPE: cDNA to mRNA														
	(vi)ORIGINAL SOURCE: (A)ORGANISM: mouse (G)CELL TYPE: liver														
20	<pre>(ix)FEATURE: (A)NAME/KEY: mat peptide (B)LOCATION: 1471 (C)IDENTIFICATION METHOD: S</pre>														
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:														
	Asn Phe Gly Arg Leu His Cys Thr Thr Ala Val Ile Arg Asn Ile Asn 1 10 15	8													
30	GAC CAA GTT CTC TTC GTT GAC AAA AGA CAG CCT GTG TTC GAG GAT ATG 9 Asp Gln Val Leu Phe Val Asp Lys Arg Gln Pro Val Phe Glu Asp Met 20 25 30	б													
	ACT GAT ATT GAT CAA AGT GCC AGT GAA CCC CAG ACC AGA CTG ATA ATA 14 Thr Asp Ile Asp Gln Ser Ala Ser Glu Pro Gln Thr Arg Leu Ile Ile 35 40 45	4													
35	TAC ATG TAC AAA GAC AGT GAA GTA AGA GGA CTG GCT GTG ACC CTC TCT 19 Tyr Met Tyr Lys Asp Ser Glu Val Arg Gly Leu Ala Val Thr Leu Ser 50 55 60	2													
	GTG AAG GAT AGT AAA ATG TCT ACC CTC TCC TGT AAG AAC AAG ATC ATT 24 Val Lys Asp Ser Lys Met Ser Thr Leu Ser Cys Lys Asn Lys Ile Ile 65 70 75 80	0													
40	TCC TTT GAG GAA ATG GAT CCA CCT GAA AAT ATT GAT GAT ATA CAA AGT 28 Ser Phe Glu Glu Met Asp Pro Pro Glu Asn Ile Asp Asp Ile Gln Ser 85 90 95	8													
	GAT CTC ATA TTC TTT CAG AAA CGT GTT CCA GGA CAC AAC AAG ATG GAG 33 Asp Leu Ile Phe Phe Gln Lys Arg Val Pro Gly His Asn Lys Met Glu 100 105 110	6													
45	TTT GAA TCT TCA CTG TAT GAA GGA CAC TTT CTT GCT TGC CAA AAG GAA 38 Phe Glu Ser Ser Leu Tyr Glu Gly His Phe Leu Ala Cys Gln Lys Glu 115 120 125	4													
50	GAT GAT GCT TTC AAA CTC ATT CTG AAA AAA AAG GAT GAA AAT GGG GAT 43 Asp Asp Ala Phe Lys Leu Ile Leu Lys Lys Lys Asp Glu Asn Gly Asp 130 135	2													
30	AAA TCT GTA ATG TTC ACT CTC ACT AAC TTA CAT CAA AGT Lys Ser Val Met Phe Thr Leu Thr Asn Leu His Gln Ser 145 150 155	1													
55	(20) INFORMATION FOR SEQ ID NO: 19:														
55	(i) SEQUENCE CHARACTERISTICS:														

(A) LENGTH: 9 amino acids (B) TYPE: amino acid (D) TCPOLOGY: linear 5 (ii) MOLECULE TYPE: peptide (v) FRAGMENT TYPE: N-terminal fragment (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19: 10 Asn Phe Gly Arg Leu His Cys Thr Thr (21) INFORMATION FOR SEQ ID NO: 20: 15 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 157 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20: Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 25 30 Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 35 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 60 30 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 35 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 115 120 125 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150 40 (22) INFORMATION FOR SEQ ID NO: 21: (i) SEQUENCE CHARACTERISTICS: 45 (A) LENGTH: 157 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21: Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 55 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile

40 45 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 50 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 110 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 115 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

(23) INFORMATION FOR SEQ ID NO: 22:

10

15

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 110 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 120 125 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 145 150

(24) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 55 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp

			20					25					3.0		
		35	Ser				40					45	Ile		
	50		Tyr			55					60	Ala			
65			Ser		70					75	Cys			_	a o
			Lys	85					9.0					95	Lys
Ser	Asp	Ile	Ile 100	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
Met	Gln	Phe 115	Glu	Ser	Ser	Ser	Tyr 120	Glu	Gly	Tyr	Phe	Leu 125	Ala	Ser	Glu
Lys	Glu 130	Arg	Asp	Leu	Phe	Lys 135	Leu	Ile	Leu	Lys	Lys 140	Glu	Asp	Glu	Leu
Gly 145	Asp	Arg	Ser	Ile	Met 150	Phe	Thr	Val	Gln	Asn 155	Glu	Asp			
(25)	TNE	CORMZ	ATTON	I FOR	ਰਸ਼ਟ	חד ו	NO.	24.							

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 45 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 50 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ser Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 105 Met Gln Phe Glu Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 115 120 125 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 130 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

- (26) INFORMATION FOR SEQ ID NO: 25:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:
- Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 55 10

5

10

15

20

25

30

35

40

45

Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 4.0 35 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ala Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 90 85 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu 120 115 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp

(27) INFORMATION FOR SEQ ID NO: 26:

5

10

15

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn 10 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp 25 Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile 40 4.5 35 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile 55 60 Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Ala Glu Asn Lys Ile 70 75 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys 85 90 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys 100 105 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Ser Glu 120 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu 135 140 130 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp 150

- (28) INFORMATION FOR SEQ ID NO: 27:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Asn Phe Gly Arg Leu His Ala Thr Thr Ala Val Ile Arg Asn Ile Asn Asp Gln Val Leu Phe Val Asp Lys Arg Gln Pro Val Phe Glu Asp Met 20 25 Thr Asp Ile Asp Gln Ser Ala Ser Glu Pro Gln Thr Arg Leu Ile Ile 35 40 Tyr Met Tyr Lys Asp Ser Glu Val Arg Gly Leu Ala Val Thr Leu Ser 55 Val Lys Asp Ser Lys Met Ser Thr Leu Ser Cys Lys Asn Lys Ile Ile 70 75 Ser Phe Glu Glu Met Asp Pro Pro Glu Asn Ile Asp Asp Ile Gln Ser 85 90 Asp Leu Ile Phe Phe Gln Lys Arg Val Pro Gly His Asn Lys Met Glu 105 Phe Glu Ser Ser Leu Tyr Glu Gly His Phe Leu Ala Cys Gln Lys Glu 115 120 125 Asp Asp Ala Phe Lys Leu Ile Leu Lys Lys Lys Asp Glu Asn Gly Asp 135 Lys Ser Val Met Phe Thr Leu Thr Asn Leu His Gln Ser 150

- (29) INFORMATION FOR SEQ ID NO: 28:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Asn Phe Gly Arg Leu His Cys Thr Thr Ala Val Ile Arg Asn Ile Asn Asp Gln Val Leu Phe Val Asp Lys Arg Gln Pro Val Phe Glu Asp Met Thr Asp Ile Asp Gln Ser Ala Ser Glu Pro Gln Thr Arg Leu Ile Ile Tyr Met Tyr Lys Asp Ser Glu Val Arg Gly Leu Ala Val Thr Leu Ser 55 Val Lys Asp Ser Lys Met Ser Thr Leu Ser Cys Lys Asn Lys Ile Ile 70 75 Ser Phe Glu Glu Met Asp Pro Pro Glu Asn Ile Asp Asp Ile Gln Ser 90 Asp Leu Ile Phe Phe Gln Lys Arg Val Pro Gly His Asn Lys Met Glu 100 105 110 Phe Glu Ser Ser Leu Tyr Glu Gly His Phe Leu Ala Ser Gln Lys Glu 115 120 125 Asp Asp Ala Phe Lys Leu Ile Leu Lys Lys Lys Asp Glu Asn Gly Asp 130 135 Lys Ser Val Met Phe Thr Leu Thr Asn Leu His Gln Ser 150

Claims

10

15

20

25

30

35

40

45

- 1. An osteoclastgenic inhibitory agent, which comprises an interleukin-18 or its functional equivalent.
- The inhibitory agent of claim 1, wherein said interleukin-18 includes the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3 as partial amino acid sequences.
 - 3. The inhibitory agent of claim 1, wherein said interleukin-18 includes the amino acid sequences of SEQ ID NO: 4

and SEQ ID NO 5 as partial amino acid sequences.

- 4. The inhibitory agent of claim 1, wherein said interleukin-18 includes the amino acid sequence of SEQ ID NO: 6.
- 5 5. The inhibitory agent of claim 1, wherein said interleukin-18 is human origin.
 - 6. The inhibitory agent of claim 1, wherein said interleukin-18 includes the amino acid sequence of SEQ ID NO: 7.
 - 7. The inhibitory agent of claim 1, which is a therapeutic agent for osteoclast-related diseases,
 - 8. The inhibitory agent of claim 1, which contains a protein, buffer, or saccharide as a stabilizer.
 - 9. The inhibitory agent of claim 1, which is in the form of a liquid, paste, or solid.
- 15. The inhibitory agent of claim 1, which contains 0.000002-100 w/w % of said interleukin-18.
 - 11. An inhibitory agent as defined in any preceding claim, for use as a pharmaceutical.
- 12. Use of an inhibitory agent as defined in any of claims 1-10 for the preparation of a medicament effective for treating and/or preventing osteoclast-related diseases.

53

10

25

30

35

40

45

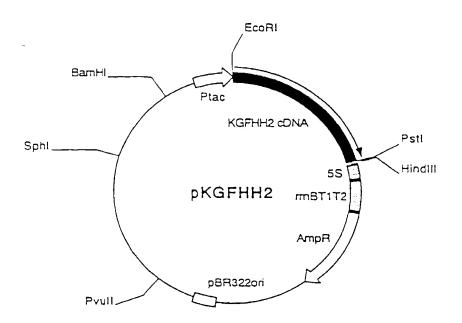


FIG. 1

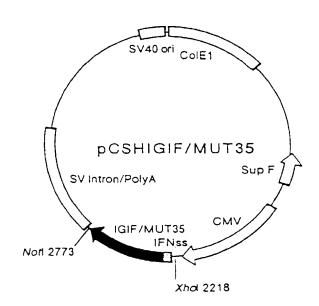


FIG. 2

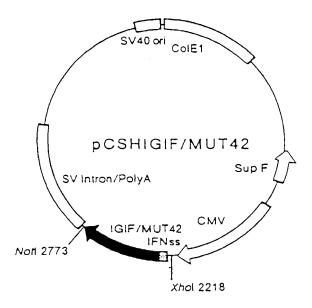


FIG. 3

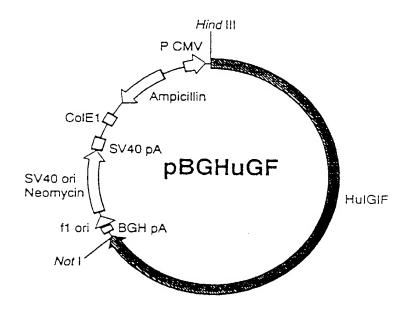


FIG. 4

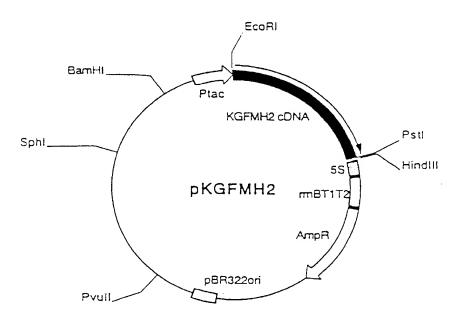


FIG. 5